

60/064,926

60/065,367

09/071,715

09/084,471

## WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISH	HED U	JNDER THE PATENT COOPERATION TREATY (PCI)
(51) International Patent Classification <sup>6</sup> :		(11) International Publication Number: WO 99/09164
C12N 15/12, C07K 14/47, C12Q 1/68, A61K 48/00, C07K 16/18, A61K 38/17	A1	(43) International Publication Date: 25 February 1999 (25.02.99)
(21) International Application Number: PCT/US  (22) International Filing Date: 14 August 1998 (		Pennsylvania Avenue N.W., Washington, DC 20004 (05).
(30) Priority Data: 60/055,784 15 August 1997 (15.08.97)	τ	(81) Designated States: AU, CA, IL, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

US

US

US

US

- (71) Applicant: ONCORMED, INC. [US/US]; 205 Perry Parkway, Gaithersburg, MD 20877 (US).
- (72) Inventors: MURPHY, Patricia, D.; 16 Stockbridge Road, Slingerlands, NY 12159-0639 (US). WHITE, Marga, B.; 8323 Sharon Drive, Frederick, MD 21704 (US). RABIN, Mark, B.; 1516 Defoe Street, Rockville, MD 20850 (US). OLSON, Sheri, J.; 2854 Yarling Court, Falls Church, VA 22042 (US). YOSHIKAWA, Matthew; 19300 Circle Gate Drive, Germantown, MD 20874 (US). JACKSON, Geoffrey, M.; 3312 Major Denton Drive, Beltsville, MD 20705 (US). ESKANDARI, Tara; 5440 Marinelli Road, Rockville, MD 20852 (US). SCHRYER, Brenda; 4240 S. Harbor Boulevard, Oxnard, CA 93035 (US). PARK, Michael; 770 Ivy League Lane, Rockville, MD 20850 (US).

- NL, PI, SE).

#### Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amen inents.

(54) Title: CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

7 November 1997 (07.11.97)

1 May 1998 (01.05.98)

22 May 1998 (22.05.98)

12 November 1997 (12.11.97)

#### (57) Abstract

Five DNA and protein sequences have been determined for the BRCA2 gene, as have been ten polymorphic sites and their rates of occurrence in the normal alleles of BRCA2. The sequences BRCA2(omi1-5) and the ten polymorphic sites will provide accuracy and reliability for genetic testing. One skilled in the art will be able to avoid misinterpretations of changes in the gene and/or protein sequence, determine the presence of a normal sequence, and of mutations of BRCA2. This invention is also related to a method of performing gene therapy with BRCA2(omil-5) coding sequences or fragments thereof. This invention is further related to protein therapy with BRCA2(omil-5) proteins or their functional equivalents.

> ATTORNEY DOCKET NUMBER: 9301-123 SERIAL NUMBER: 09/724,538

REFERENCE: CY

#### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel ·	MR	Mauritania	UG	Ugunda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	It <b>aly</b>	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
KE	Estonia	LR	Liberia	SG	Singapore		

#### CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

This is an U.S. utility patent application based on U.S. Provisional Application Serial Nos. 60/055,784 filed on August 15, 1997, 60/064,926 filed on November 7, 1997, and 60/065,367 filed on November 12, 1997.

#### FIELD OF THE INVENTION

This invention relates to a gene which has been associated with breast cancer where the gene is found to be mutated. More specifically, this invention relates to five unique coding sequences of BRCA2 gene BRCA2<sup>(omi1)</sup>, BRCA2<sup>(omi2)</sup>, BRCA2<sup>(omi3)</sup>, BRCA2<sup>(omi4)</sup>, and BRCA2<sup>(omi5)</sup> identified in human subjects which define five novel haplotypes.

#### BACKGROUND OF THE INVENTION

It has been estimated that about 5-10% of breast cancer is inherited (Rowell, S., et al., American Journal of Human Genetics 55:861-865 (1994)). The first gene associated with both breast and ovarian cancer was cloned in 1994 from chromosome 17 by Miki, Y., et al., Science 266:66-71 (1994). A second high-risk breast cancer conferring gene was located on chromosome 13 in 1994 (Wooster, R., et al., Science 265:2088-2090) and subsequently cloned in 1995 (Wooster, R., et al., Nature 378:789-792). Mutations in this "tumor suppressor" gene are thought to account for roughly 35% of inherited breast cancer and 80-90% of families with male breast cancer.

Locating one or more mutations in the BRCA2 region of chromosome 13 provides a promising approach to reducing the high incidence and mortality associated with breast cancer through the early detection of women and men at high risk. These individuals, once identified, can be targeted for more aggressive prevention programs. Screening is carried out by a variety of methods which include karyotyping, probe binding and DNA sequencing.

In DNA sequencing technology, genomic DNA is extracted from whole blood and the coding regions of the BRCA2 gene are amplified. Each of the coding regions may be sequenced completely and the results are compared to the normal DNA sequence of the gene. Alternatively, the coding sequence of the sample gene may be compared to a panel of known mutations or other

30

25

5

10

15

20

5

10

15

20

25

30

screening procedure before completely sequencing the gene and comparing it to a normal sequence of the gene.

The BRCA2 gene is divided into 27 separate exons. Exon 1 is noncoding, in that it is not part of the final functional BRCA2 protein product. The BRCA2 coding region spans roughly 10433 base pairs (bp) over 70 kb. Each exon consists of 100-600 bp, except for exons 10, 11 and 27. The full length mRNA is 11-12 kb. To sequence the coding region of the BRCA2 gene, each exon is amplified separately and the resulting PCR products are sequenced in the forward and reverse directions. Because exons 10, 11, and 27 are so large, we have divided them into three, twenty-one, and two overlapping PCR fragments (respectively) of approximately 250-625 bp each (segments "A" through "C" of exon 10, "A" through "U" of exon 11, and "A" through "B" of exon 27).

Many mutations and normal polymorphisms have already been reported in the BRCA2 gene. A world wide web site has been built to facilitate the detection and characterization of alterations in breast cancer susceptibility genes. Such mutations in BRCA2 can be accessed through the Breast Cancer Information Core (BIC) at <a href="http://www.nhgri.nih.gov/Intramural\_research/Lab\_transfer/Bic">http://www.nhgri.nih.gov/Intramural\_research/Lab\_transfer/Bic</a>. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995). The information on BRCA2 was added in February, 1996.

The genetics of Breast Cancer Syndrome is autosomal dominant with reduced penetrance. In simple terms, this means that the syndrome runs through families: (1) both sexes can be carriers (mostly women get the disease but men can both pass it on and occasionally get breast cancer); (2) most generations will likely have breast cancer; (3) occasionally women carriers either die young before they have the time to manifest disease (and yet have offspring who get it) or they never develop breast or ovarian cancer and die of old age (the latter people are said to have "reduced penetrance" because they never develop cancer). Pedigree analysis and genetic counseling is absolutely essential to the proper workup of a family prior to any lab work.

Until now, the only sources of genomic sequence information for BRCA2 were GenBank (Accession Number U43746), or through the Breast Information Core (BIC) database on the Internet which requires membership in the BIC consortium. However, based upon the disclosure of this patent application, in neither GenBank nor BIC were the sequences identified and listed entirely accurate. There is a need in the art to correct these mistakes which otherwise may lead to misinterpretation of the sequence data from the patient as abnormal when it was not, or vice versa.

In addition, there is a need in the art to have available a functional allele profile which represents the most likely BRCA2 sequences to be found in the majority of the normal population. This functional allele profile is based upon frequent polymorphisms and the correct backbone sequence. The knowledge of several common normal haplotypes will make it possible for true mutations to be easily identified or differentiated from polymorphisms. Identification of mutations of the BRCA2 gene and protein would allow more widespread diagnostic screening for hereditary breast cancer than is currently possible.

The use of these common normal haplotypes, in addition to the previously published BRCA2 sequence, will reduce the likelihood of misinterpreting a "sequence variation" found in the normal population with a pathologic "mutation" (i.e. causes disease in the individual or puts the individual at a high risk of developing the disease). With large interest in breast cancer predisposition testing, misinterpretation is particularly worrisome. People who already have breast cancer are asking the clinical question: "is my disease caused by a heritable genetic mutation?" The relatives of the those with breast cancer are asking the question: "Am I also a carrier of the mutation my relative has? Thus, is my risk increased, and should I undergo a more aggressive surveillance program?"

#### 30 SUMMARY OF THE INVENTION

5

10

15

20

25

The present invention is based on the discovery of the correct genomic BRCA2 sequence and five novel sequence haplotypes found in normal human subjects of the BRCA2 gene.

5

10

15

20

25

30

It is an object of this invention to provide the correct intronic/exonic sequence of the BRCA2 gene.

It is another object of this invention to provide five unique haplotype sequences of the BRCA2 gene in normal individuals which do not correspond to increased cancer susceptibility.

It is another object of this invention to sequence a BRCA2 gene or a portion thereof and compare it to the five haplotype sequences to determine whether a sequence variation noted represents a polymorphism or a potentially harmful mutation.

It is another object of this invention to provide a list of the pairs which occur at each of ten polymorphic points in the BRCA2 gene.

It is another object of this invention to provide the rates of occurrence for the polymorphisms at codons 289, 372, 455, 743, 894, 991, 1132, 1269, 2414, and 2951 in the BRCA2 gene.

It is another object of this invention to provide a method wherein all exons of BRCA2 gene or parts thereof, are amplified with one or more oligonucleotide primers.

It is another object of this invention to provide a method of identifying a individual who carries no mutation(s) of the BRCA2 gene and is therefore at no increased risk or susceptibility to breast or ovarian cancer based on a finding that the individual does not carry an abnormal BRCA2 genes.

It is another object of this invention to provide a method of identifying a mutation in BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer.

It is another object of this invention to provide five novel BRCA2 protein sequences derived from five BRCA2 haplotype sequences.

It is another object of the invention to encompass prokaryotic or eukaryotic host cells comprising an expression vector having a DNA sequence that encodes for all or a fragment of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof, or a functional equivalent thereof.

It is another object of the invention to encompass an anti-BRCA2 protein antibody using all of fragments of the five novel BRCA2 protein

sequences, a BRCA2 polypeptide thereof or a functional equivalent thereof as an immunogen.

There is a need in the art for cDNA sequences of the BRCA2 gene and for the protein sequences of BRCA2 gene from normal individuals who are not at risk for increased susceptibility for cancer. In order to determine whether a sample from a patient suspected of containing a BRCA2 mutation actually has the mutation, the patient's BRCA2 DNA and/or amino acid sequence need to be compared to all known normal BRCA2 sequences. Failure to compare the sequence obtained to all naturally occurring normal sequences may result in reporting a sample as containing a potentially harmful mutation when it is a polymorphism without clinical significance.

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene with no coding sequence mutations, who therefore cannot be said to have an increased genetic susceptibility to breast or ovarian cancer from their BRCA2 genes;
- avoiding misinterpretation of normal polymorphisms found in the BRCA2 gene;
- c) determining the presence of a previously unknown mutation in the BRCA2 gene;
  - identifying a mutation in exon 11 of BRCA2 which indicates a predisposition or higher susceptibility to ovarian cancer than

breast

25

5

10

15

- cancer (i.e., resides in the putative "ovarian cancer cluster" region);
  - e) probing a human sample of the BRCA2 gene by allele to determine the presence or absence of either polymorphic alleles or mutations;
- of) performing gene therapy with the correct BRCA2 gene sequence.
  - g) performing protein replacement therapy with the correct BRCA 2 protein sequence or a functional equivalent thereof.

# BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 shows the GenBank genomic sequence of BRCA2 (Accession Number U43746). The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Incorrect exonic sequences at exons 5 and 16 are shown with boldface type.

FIGURE 2 shows the corrected genomic sequence of BRCA2. The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Corrected intronic and exonic sequences at exons 5, 11 and 15 are shown with boldface type.

FIGURE 3 shows the alternative alleles at polymorphic sites along a chromosome which can be represented as a unit or "haplotype" within a gene such as BRCA2. The haplotype that is in GenBank (GB) is shown with light shading. Five additional haplotypes are shown in FIGURE 3

(encompassing the alternative alleles found at nucleotide sites 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470 and 9079). BRCA2 (omi-1), BRCA2 (omi-2), BRCA2 (omi-3), BRCA2 (omi-4), and BRCA2 (omi-5) are represented with mixed dark and light shading (numbers 2, 4, 6, 8 and 10 from left to right). In total, 5 of 10 haplotypes along the BRCA2 gene are unique.

20

25

30

15

5

10

# DETAILED DESCRIPTION OF THE INVENTION

#### **DEFINITIONS**

The following definitions are provided for the purpose of understanding this invention.

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast, ovarian and pancreatic cancer in women and also breast, prostate and pancreatic cancer in men. BRCA2 is associated with genetic susceptibility to breast, ovarian and pancreatic cancer. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian, prostate, and pancreatic cancers in men and women.

"Coding sequence" refers to those portions of a gene which, taken together, code for a peptide (protein), or which nucleic acid itself has function.

"Protein" or "peptide" refers to a sequence of amino acids which has function.

"BRCA2<sup>(omi)</sup>" refers to the genomic BRCA2 sequence disclosed in Genbank (Accession Number U43746) wherein,

5

10

15

20

25

30

(1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5; and

(2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon.

"BRCA2<sup>(omi 1-5)</sup>" refers to five unique DNA sequences of the BRCA2 gene and their introns (particularly the slice sites adjacent to the exons). These sequences were found by end to end sequencing of the BRCA2 gene from 5 individuals randomly drawn from the population and who were documented to have no family history of breast or ovarian cancer. The sequenced exons were found not to contain any truncating mutations. In all cases the change of a nucleic acid at a polymorphic site lead to a codon change and a change of amino acid from the previously published standard in GenBank (see TABLE III). In some cases the frequency of occurrence of a nucleic acid change was found to differ from the published frequency or was newly determined. These sequence variations are believed to be alleles whose haplotypes do not indicate an increased risk for cancer.

"Normal DNA sequence" also called "normal gene sequence" refere to a nucleic acid sequence, the nucleic acid of which are known to occur at their respective positions with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which itself has normal function.

"Normal Protein Sequence" refers to the protein sequence, the amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein.

"Normal Sequence" refers to the nucleic acid or protein sequence, the nucleic or amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which nucleic acid itself has a normal function.

"Haplotype" refers to a series of specific alleles within a gene along a chromosome.

"Functional allele profile" refers a list of those alleles in the normal population which have the funll function.

5

10

15

20

25

30

"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence coding for a non-functional protein or a protein with substantially reduced or altered function.

"Polymorphism" refers to a base change in a DNA sequence which is not associated with known pathology.

"Primer" refers to a sequence comprising about 15 or more nucleotides having a sequence complementary to the BRCA2 gene. Other primers which can be used for primer hybridization will be known or readily ascertainable to those skilled in the art.

"Substantially complementary to" refers to primer sequences which hybridize to the sequences provided under stringent conditions and/or sequences having sufficient homology with BRCA2 sequences, such that the allele specific oligonucleotide primers hybridize to the BRCA2 sequences to which they are complimentary.

"Isolated nucleic acids" refers to nucleic acids substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they may be associated. Such association is typically either in cellular material or in a synthesis medium.

"Biological sample" or "body sample" refers to a sample containing DNA oatained from a biological source. The sample may be from a living, dead or even archeological source from a variety of tissues and cells. Examples include body fluid (e.g. blood (ieukocytes), urine (epithelial cells), saliva, breast milk, menstrual flow, cervical and vaginal secretions, etc.), skin, hair roots/follicle, mucus membrane (e.g. buccal or tongue cell scrapings), cervicovaginal cells (from PAP smear, etc.), lymphatic tissue, internal tissue (normal or tumor).

"Vector" refers to any polynucleotide which is capable of self replication or inducing integration into a self-replicating polynucleotide.

Examples include polynucleotides containing an origin or replication or an

integration site. Vectors may be intergrated into the host cell's chromosome or form an autonomously replicating unit.

"A tumor growth inhibitor" refers to a molecule such as, all or a fragment of BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent thereof that is effective for preventing the formation of, reducing, or eliminating a transformed or malignant phenotype of breast or ovarian cancer cells.

"A BRCA2 polypeptide" refers to a BRCA2 polypeptide either directly derived from the BRCA2 protein, or homologous to the BRCA2 protein, or a fusion protein consisting of all or fragments of the BRCA2 protein and polypeptides.

"A functional equivalent" refers to a molecule including an unnatural BRCA2 polypeptide, a drug or a natural product which retains substantial biological activity as the native BRCA2 protein. The activity and function of BRCA2 protein may include transactivation, granin, DNA repair, among others.

"A target polynucleotide" refers to the nucleic acid sequence of interest, for example, the BRCA2 encoding polynucleotide. Other primers which can be used for primer hybridization will be known or readily ascertainable to those of skill in the art.

The invention in several of its embodiments includes: an isolated DNA sequence of the BRCA2 coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, and 12, a protein sequence of the BRCA2 protein as set forth in SEQ ID NO:5, 7, 9, 11, 13, a method of identifying individuals having a normal BRCA2 gene with no increased risk for breast and ovarian cancer, a method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, a method of performing gene therapy to prevent or treat a tumor, a method of protein replacement therapy to prevent or treat a tumor, a diagnostic reagent comprising all or fragments of the disclosed BRCA2 cDNA and protein sequences.

SEQUENCING

5

10

15

20

25

30

5

10

15

20

25

30

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid, providing it contains, or is suspected of containing, the specific nucleic acid sequence containing a polymorphic or a mutant allele. Thus, the process may amplify, for example, DNA or RNA, including mRNA and cDNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of nucleic acids may also be employed, or the nucleic acids produced in a previous method such as an amplification reaction using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified, i.e., the polymorphic and/or the mutant allele, may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. A variety of amplification techniques may be used such as ligating the DNA sample or fragments thereof to a vector capable of replication or incorporation into a replicating system thereby increasing the number of copies of DNA suspected of containing at least a portion of the BRCA2 gene. Amplification techniques include so called "shot gun cloning". It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as contained in whole human DNA.

It should be noted that one need not sequence the entire coding region or even an entire DNA fragment in order to determine whether or not a mutation is present. For example, when a mutation is known in one family member, it is sufficient to determine the sequence at only the mutation site by sequencing or by other mutation detection systems such as ASO when testing other family members.

DNA utilized herein may be extracted from a body sample, such as blood, tissue material and other biological sample by a variety of techniques such as that described by Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, 1982). If the extracted sample is impure, it may be treated before amplification with an amount of a reagent

5

10

15

20

25

30

effective to open the cells, and to expose and/or separate the strand(s) of the nucleic acid(s). This lysing and nucleic acid denaturing step to expose and separate the strands will allow amplification to occur much more readily.

For amplification by cloning, the isolated DNA may be cleaved into fragments by a restriction endonuclease or by shearing by passing the DNA containing mixture through a 25 gauge needle from a syringe to prepare 1-1.5 kb fragments. The fragments are then ligated to a cleaved vector (virus, plasmid, transposon, cosmid etc.) and then the recombinant vector so formed is then replicated in a manner typical for that vector.

For a PCR amplification, the deoxyribonucleotide triphosphates dATP. dCTP, dGTP, and dTTP are added to the synthesis mixture, either separately or together with the primers, in adequate amounts and the resulting solution is heated to about 90°-100°C from about 1 to 10 minutes. preferably from 1 to 4 minutes. After this heating period, the solution is allowed to cool, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for effecting the primer extension reaction (called herein "agent for polymerization"), and the reaction is allowed to occur under conditions known in the art. The agent for polymerization may also be added together with the other reagents if it is heat stable. This synthesis (or amplification) reaction may occur at room temperature up to a temperature above which the agent for polymerization no longer functions. Thus, for example, if DNA polymerase is used as the agent, the temperature is generally no greater than about 40°C. Most conveniently the reaction occurs at room temperature. When using thermostable DNA polymerase such as Taq, higher temperature may be used.

The allele specific oligonucleotide primers are useful in determining whether a subject is at risk of having breast or ovarian cancer, and also useful for characterizing a tumor. Primers direct amplification of a target polynucleotide prior to sequencing. These unique BRCA2 oligonucleotide primers for exons 2-27 shown in TABLE II were designed and produced specifically to optimize amplification of portions of BRCA2 which are to be sequenced.

The primers used to carry out this invention embrace oligonucleotides of sufficient length and appropriate sequence to provide initiation of polymerization. Environmental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization, such as DNA polymerase, and a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The oligonucleotide primer typically contains 18-28 bp plus in some cases an M13 "tail" for convenience.

Primers used to carry out this invention are designed to be substantially complementary to each strand of the genomic locus to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5' and 3' sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

Oligonucleotide primers of the invention are employed in the amplification process which is an enzymatic chain reaction that produces exponential quantities of polymorphic locus relative to the number of reaction steps involved. Typically, one primer is complementary to the negative (-) strand of the polymorphic locus and the other is complementary to the positive (+) strand. Accealing the primers to denatured nucleic acid followed by extension with an enzyme, such as the large fragment of DNA polymerase I (Klenow) and nucleotides, results in newly synthesized + and strands containing the target polymorphic locus sequence. Because these newly synthesized sequences are also templates, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the region (i.e., the target polymorphic locus sequence) defined by the primers. The product of the chain reaction is a discreet nucleic acid

duplex with termini corresponding to the ends of the specific primers employed.

5

10

15

20

25

30

The oligonucleotide primers of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al., Tetrahedron Letters, 22:1859-1862, 1981. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The agent for polymerization may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase, polymerase muteins, reverse transcriptase, other enzymes, including heat-stable enzymes (*i.e.*, those enzymes which perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation), such as *Taq* polymerase. Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each polymorphic locus nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths.

The newly synthesized strand and its complementary nucleic acid strand will form a double-stranded molecule under hybridizing conditions described above and this hybrid is used in subsequent steps of the process. In the next step, the newly synthesized double-stranded molecule is subjected to denaturing conditions using any of the procedures described above to provide single-stranded molecules.

The steps of denaturing, annealing, and extension product synthesis can be repeated as often as needed to amplify the target polymorphic locus nucleic acid sequence to the extent necessary for detection. The amount of the specific nucleic acid sequence produced will accumulate in an

5

10

15

20

25

30

exponential fashion. Amplification is described in *PCR. A Practical Approach*, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, 1992.

The amplification products may be detected by Southern blots analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a very low level of the nucleic acid sequence of the polymorphic locus is amplified, and analyzed via a Southern blotting technique or similarly, using dot blot analysis. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. Alternatively, probes used to detect the amplified products can be directly or indirectly detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme. Those of ordinary skill in the art will know of other suitable labels for binding to the probe, or will be able to ascertain such, using routine experimentation.

Sequences amplified by the methods of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence such as PCR, oligomer restriction (Saiki, et.al., Bio/Technology, 3:1008-1012, 1985), allele-specific oligonucleotide (ASO) probe analysis (Conner, et al., Proc. Natl. Acad. Sci. U.S.A., 80:278, 1983), oligonucleotide ligation assays (OLAs) (Landgren, et al., Science, 241:1007, 1988), and the like. Molecular techniques for DNA analysis have been reviewed (Landgren, et al., Science, 242:229-237, 1988).

Preferably, the method of amplifying is by PCR, as described herein and as is commonly used by those of ordinary skill in the art. Alternative methods of amplification have been described and can also be employed as long as the BRCA2 locus amplified by PCR using primers of the invention is similarly amplified by the alternative means. Such alternative amplification systems include but are not limited to self-sustained sequence replication, which begins with a short sequence of RNA of interest and a T7 promoter. Reverse transcriptase copies the RNA into cDNA and degrades the RNA,

followed by reverse transcriptase polymerizing a second strand of DNA. Another nucleic acid amplification technique is nucleic acid sequence-based amplification (NASBA) which uses reverse transcription and T7 RNA polymerase and incorporates two primers to target its cycling scheme. NASBA can begin with either DNA or RNA and finish with either, and 5 amplifies to 10<sup>8</sup> copies within 60 to 90 minutes. Alternatively, nucleic acid can be amplified by ligation activated transcription (LAT). LAT works from a single-stranded template with a single primer that is partially single-stranded and partially double-stranded. Amplification is initiated by ligating a cDNA to the promoter oligonucleotide and within a few hours, and amplification is 108 10 to 10<sup>9</sup> fold. Another amplification system useful in the method of the invention is the Qβ Replicase System. The Qβ replicase system can be utilized by attaching an RNA sequence called MDV-1 to RNA complementary to a DNA sequence of interest. Upon mixing with a sample, the hybrid RNA finds its complement among the specimen's mRNAs and binds, activating 15 the replicase to copy the tag-along sequence of interest. Another nucleic acid amplification technique, ligase chain reaction (LCR), works by using two differently labeled halves of a sequence of interest which are covalently bonded by ligase in the presence of the contiguous sequence in a sample, forming a new target. The repair chain reaction (RCR) nucleic acid 20 amplification technique uses two complementary and target-specific oligonucleotide probe pairs, thermostable polymerase and ligase, and DNA nucleotides to geometrically amplify targeted sequences. A 2-base gap separates the oligonucleotide probe pairs, and the RCR fills and joins the 25 gap, mimicking normal DNA repair. Nucleic acid amplification by strand displacement activation (SDA) utilizes a short primer containing a recognition site for hincl with short overhang on the 5' end which binds to target DNA. A DNA polymerase fills in the part of the primer opposite the overhang with sulfur-containing adenine analogs. Hincll is added but only cuts the 30 unmodified DNA strand. A DNA polymerase that lacks 5' exonuclease activity enters at the site of the nick and begins to polymerize, displacing the initial primer strand cownstream and building a new one which serves as

5

10

15

20

25

30

more primer. SDA produces greater than 10<sup>7</sup>-fold amplification in 2 hours at 37°C. Unlike PCR and LCR, SDA does not require instrumented Temperature cycling.

Another method is a process for amplifying nucleic acid sequences from a DNA or RNA template which may be purified or may exist in a mixture of nucleic acids. The resulting nucleic acid sequences may be exact copies of the template, or may be modified. The process has advantages over PCR in that it increases the fidelity of copying a specific nucleic acid sequence, and it allows one to more efficiently detect a particular point mutation in a single assay. A target nucleic acid is amplified enzymatically while avoiding strand displacement. Three primers are used. A first primer is complementary to the first end of the target. A second primer is complementary to the second end of the target. A third primer which is similar to the first end of the target and which is substantially complementary to at least a portion of the first primer such that when the third primer is hybridized to the first primer, the position of the third primer complementary to the base at the 5' end of the first primer contains a modification which substantially avoids strand displacement. This method is detailed in U.S. Patent 5,593,840 to Bhatnagar et al. 1997, incorporated herein by reference.

Finally, recent application of DNA chips or microarray technology where DNA or oligonucleotides are immobilized on small solid support may also be used to rapidly sequence sample BRCA2 gene and analyze its expression. Typically, high density arrays of DNA fragment are fabricated on glass or nylon substrates by *in situ* light-directed combinatorial synthesis or by conventional synthesis followed by immobilization (Fodor *et al.* U.S. patent No. 5,445,934). Sample DNA or RNA may be amplified by PCR, labeled with a fluorescent tag, and hybridized to the microarray. Examples of this technology are provided in U.S. Patents 5,510, 270, U.S. 5,547,839, incorporated herein by reference.

All exonic and adjacent intronic sequences of the BRCA2 gene were obtained by end to end sequencing of five normal subjects in the manner described above followed by analysis of the data obtained. The data obtained provided us with the opportunity to establish the correct

intronic/exonic structure of the BRCA2 gene. In addition, we evaluated six previously published normal polymorphisms (1342, 2457, 3199, 3624, 4035, and 7470) for correctness and frequency in the population, and to identify four additional polymorphisms not previously characterized (1093, 1593, 2908, and 9079).

#### **GENE THERAPY**

5

10

15

20

25

30

The polynucleotide(s) which result from either sense or antisense transcription of any exon or the entire coding sequence or fragments of BRCA2 gene may be used for gene therapy. A variety of methods are known for gene transfer, any of which might be available for use.

Direct injection of Recombinant DNA in vivo:

- 1. Direct injection of "naked" DNA directly with a syringe and needle into a specific tissue, infused through a vascular bed, or transferred through a catheter into endothelial cells.
- 2. Direct injection of DNA that is contained in artificially generated lipid vesicles or other encapsulating vehicles.
- 3. Direct injection of DNA conjugated to a target receptor structure, such as a diptheria toxin, an antibody or other suitable receptor.
- 4. Direct injection by particle bombardment. For example, the DNA may be coated onto gold particles and shot into the cells.

#### **Human Artificial Chromosomes**

The gene delivery approach involves the use of human chromosomes that have been stripped down to contain only the essential components for replication and the genes desired for transfer.

#### **keceptor-Mediated Gene Transfer**

DNA is linked to a targeting molecule that will bind to specific cell-surface receptors, inducing endocytosis and transfer of the DNA into mammalian cells. One such technique uses poly-L-lysine to link asialoglycoprotein to DNA. An adenovirus is also added to the complex to disrupt the lysosomes and thus allow the DNA to avoid degradation and move to the nucleus. Infusion of these particles intravenously has resulted in gene transfer into hepatocytes.

#### RECOMBINANT VIRUS VECTORS

Several vectors may be used in gene therapy. Among them are the Moloney Murine Leukemia Virus (MoMLV) Vectors, the adenovirus vectors, the Adeno-Associated Virus (AAV) vectors, the herpes simplex virus (HSV) vectors, the poxvirus vectors, the retrovirus vectors, and human immunodeficiency virus (HIV) vectors.

#### GENE REPLACEMENT AND REPAIR

The ideal genetic manipulation for treatment of a genetic disease would be the actual replacement of the defective gene with a normal copy of the gene. Homologous recombination is the term used for switching out a section of DNA and replacing it with a new piece. By this technique, the defective gene may be replaced with a normal gene which expresses a functioning BRCA2 tumor growth inhibitor protein.

A complete description of gene therapy can also be found in "Gene Therapy A Primer For Physicians" 2d Ed. by Kenneth W. Culver, M.D. Publ. Mary Ann Liebert Inc. (1996). Two Gene Therapy Protocols for BRCA1 gene have been approved by the Recombinant DNA Advisory Committee for Jeffrey T. Holt et al. They are listed as 9602-148, and 9603-149 and are available from the NIH. Protocols for BRCA2 gene therapy may be similarly employed. The isolated BRCA2 gene may be synthesized or constructed from amplification products and inserted into a vector such as the LXSN vector.

25

30

5

10

15

20

#### A BRCA2 POLYPEPTIDE OR ITS FUNCTIONAL EQUIVALENT

The growth of breast and ovarian cancer may be arrested or prevented by directly increasing the BRCA2 protein level where inadequate functional BRCA2 activity is responsible for breast and ovarian cancer. The cDNA and amino acid sequences of five novel BRCA2 haplotypes are disclosed herein (SEQ ID No:4-13). All or a fragment of BRCA2 protein may be used in therapeutic or prophylactic treatment of breast and ovarian cancer. Such a fragment may have a similar biological function as the native BRCA2 protein

5

10

15

20

25

30

or may have a desired biological function as specified below. BRCA2 polypeptides or their functional equivalents including homologous and modified polypeptide sequences are also within the scope of the present invention. Changes in the native sequence may be advantageous in producing or using the BRCA2 derived polypeptides or functional equivalents suitable for therapeutic or prophylactic treatment of breast and ovarian cancer. For example, these changes may be desirable for producing resistance against *in vivo* proteolytic cleavage, for facilitating transportation and delivery of therapeutic reagents, for localizing and compartmentalizing tumor suppressing agents, or for expression, isolating and purifying the target species.

There are a variety of methods to produce an active BRCA2 polypeptide or a functional equivalent as a tumor growth inhibitor. For example, one or more amino acids may be substituted, deleted, or inserted using methods well known in the art (Maniatis et al., 1982). Considerations of polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphiphathic nature of the amino acids play an important role in designing homologous polypeptide changes suitable for the intended treatment. In particular, conservative amino acid substitution using amino acids that are related in side-chain structure and charge may be employed to preserve the chemical and biological property. A homologous polyeptide typically contains at least 70% homology to the native sequence. Unnatural forms of the polypeptide may also be incorporated so long as the modification retains substantial biological activity. These unnatural polypeptides typically include structural mimics and chemical medications, which have similar threedimensional structures as the active regions of the native BRCA2 protein. For example, these modifications may include terminal D-amino acids, cyclic peptides, unnatural amino acids side chains, pseudopeptide bonds, Nterminal acetylation, glycosylation, and biotinylation, etc. These unnatural forms of polypeptide may have a desired biological function, for example, they may be particularly robust in the presence of cellular or serum proteases and exopeptidase. An effective BRCA2 polypeptide or a functional equivalent may also be recognized by the reduction of the native

5

10

15

20

25

30

BRCA2 protein. Regions of the BRCA2 protein may be systematically deleted to identify which regions are essential for tumor growth inhibitor activity. These smaller fragments of BRCA2 protein may then be subjected to structural and functional modification to derive therapeutically or prophylactically effective regiments. Finally, drugs, natural products or small molecules may be screened or synthesized to mimic the function of the BRCA2 protein. Typically, the active species retain the essential threedimensional shape and chemical reactivity, and therefore retain the desired aspects of the biological activity of the native BRCA2 protein. The activity and function of BRCA2 may include transactivation, granin, DNA repair among others. Functions of BRCA2 protein are also reviewed in Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998) and Zhang et al., Cell 92:433-436 (1998). It will be apparent to one skilled in the art that a BRCA2 polypeptide or a functional equivalent may be selected because such polypeptide or functional equivalent possesses similar biological activity as the native BRCA2 protein.

# EXPRESSION OF THE BRCA2 PROTEIN AND POLYPEPTIDE IN HOST CELLS

All or fragments of the BRCA2 protein and polypeptide may be produced by host cells that are capable of directing the replication and the expression of foreign genes. Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic cells, which contain an expression vector comprising all or a fragment of the BRCA2 cDNA sequence (SEQ. ID No: 4, 6, 8, 10, or 12) operatively linked to one or more regulatory sequences to produce the intended BRCA2 protein or polypeptide. Prokaryotes may include gram negative or gram positive organisms, for example *E. coli* or *Bacillus* strains. Suitable eukaryotic host cells may include yeast, virus, and mamalian systems. For example, Sf9 insect cells and human cell lines, such as COS, MCF7, HeLa, 293T, HBL100, SW480, and HCT116 cells.

A broad variety of suitable expression vectors are available in the art.

An expression vector typically contains an origin of replication, a promoter, a phenotypic selection gene (antibiotic resistance or autotrophic requirement),

5

10

15

20

25

30

and a DNA sequence coding for all or fragments of the BRCA2 protein. The expression vectors may also include other operatively linked regulatory DNA sequences known in the art, for example, stability leader sequences, secretory leader sequences, restriction enzyme cleavage sequences, polyadenylation sequences, and termination sequences, among others. The essential and regulatory elements of the expression vector must be compatible with the intended host cell. Suitable expression vectors containing the desired coding and control regions may be constructed using standard recombinant DNA techniques known in the art, many of which are described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). For example, suitable origins of replication may include Col E1, SV4O viral and M13 origins of replication. Suitable promoters may be constitutive or inducible, for example, tac promoter, lac Z promoter, SV40 promoter, MMTV promoter, and LXSN promoter. Examples of selectable markers include neomycin, ampicillin, and hygromycin resistance and the like. Many suitable prokaryotic, viral and mammalian expression vectors may be obtained commercially, for example, from Invitrogen Corp., San Diego, CA or from Clontech, Palo Alto, CA. It may be desirable that the BRCA2 protein or polypeptide is produced as a fusion protein to enhance the expression in selected host cells, to detect the expression in transfected cells, or to simplify the purification process. Suitable fusion partners for the BRCA2 protein or polypeptide are well known in the art and may include  $\beta$ -galactosidase, glutathione-S-transferase, and poly-histidine tag.

Expression vectors may be introduced into host cells by various methods known in the art. The transformation procedure used depends upon the host to be transformed. Methods for introduction of vectors into host cells may include calcium phosphate precipitation, electrosporation, dextranmediated transfection, liposome encapsulation, nucleus microinjection, and viral or phage infection, among others.

Once an expression vector has been introduced into a suitable host cell, the host cell may be cultured under conditions permitting expression of large amounts of the BRCA2 protein or polypeptide. The expression product

5

10

15

20

25

30

may be identified by many approaches well known in the art, for example, sequencing after PCR-based amplification, hybridization using probes complementary to the desired DNA sequence, the presence or absence of marker gene functions such as enzyme activity or antibiotic resistance, the level of mRNA production encoding the intended sequence, immunological detection of a gene product using monoclonal and polyclonal antibodies, such as Western blotting or ELISA. The BRCA2 protein or polypeptides produced in this manner may then be isolated following cell lysis and purified using various protein purification techniques known in the art, for example, ion exchange chromatography, gel filtration chamatography and immunoaffinity chromatography.

It is generally preferred that whenever possible, longer fragments of BRCA2 protein or polypeptide are used, particularly to include the desired functional domains of BRCA2 protein. Expression of shorter fragments of DNA may be useful in generating BRCA2 derived immunogen for the production of anti-BRCA2 antibodies. It should, of course, be understood that not all expression vectors, DNA regulatory sequences or host cells will function equally well to express the BRCA2 protein or polypeptides of the present invention. However, one of ordinary skill in the art may make a selection among expression vectors, DNA regulatory sequences, host cells, and codon usage in order to optimize expression using known technology in the art without undue experimentation. Studies of BRCA2 protein function and examples of genetic manipulation of BRCA2 protein are summarized in two recent review articles, Bertwistle and Ashworth, *Curr. Opin. Genet. Dev.* 8(1): 14-20 (1998) and Zhang *et al.*, *Cell* 92:433-436 (1998).

### IN VITRO SYNTHESIS AND CHEMICAL SYNTHESIS

Although it is preferred that fragments of the BRCA2 protein or polypeptides be obtained by overexpression in prokaryotic or eukaryotic host cells, the BRCA2 polypeptides or their functional equivalents may also be obtained by *in vitro* translation or synthetic means by methods known to those of ordinary skill in the art. For example, *in vitro* translation may employ an mRNA encoded by a DNA sequence coding for fragments of the BRCA2

protein or polypeptides. Chemical synthesis methodology such as solid phase synthesis may be used to synthesize a BRCA2 polypeptide structural mimic and chemically modified analogs thereof. The polypeptides or the modifications and mimic thereof produced in this manner may then be isolated and purified using various purification techniques, such as chromatographic procedures including ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

#### PROTEIN REPLACEMENT THERAPY

10

15

20

25

30

The tumor suppressing function of BRCA2 suggests that various BRCA2 protein targeted therapies may be utilized in treating and preventing tumors in breast and ovarian cancer. The present invention therefore includes therapeutic and prophylactic treatment of breast and ovarian cancer using therapeutic pharmaceutical compositions containing the BRCA2 protein, polypeptides, or their functional equivalents. For example, protein replacement therapy may involve directly administering the BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent in a pharmaceutically effective carrier. Alternatively, protein replacement therapy may utilize tumor antigen specific antibody fused to fragments of the BRCA2 protein, a polypeptide, or a functional equivalent to deliver anti-cancer regiments specifically to the tumor cells.

To prepare the pharmaceutical compositions of the present invention, an active BRCA2 protein, a BRCA2 polypeptide, or its functional equivalent is combined with a pharmaceutical carrier selected and prepared according to conventional pharmaceutical compounding techniques. A suitable amount of the composition may be administered locally to the site of a tumor or systemically to arrest the proliferation of tumor cells. The methods for administration, may include parenteral, oral, or intravenous, among others according to established protocols in the art.

Pharmaceutically acceptable solid or liquid carriers or components which may be added to enhance or stabilize the composition, or to facilitate preparation of the composition include, without limitation, syrup, water, isotonic solution, 5 % glucose in water or buffered sodium or ammonium

acetate solution, oils, glycerin, alcohols, flavoring agents, preservatives, coloring agents, starches, sugars, diluents, granulating agents, lubricants, binders, and sustained release materials. The dosage at which the therapeutic compositions are administered may vary within a wide range and depends on various factors, such as the stage of cancer progression, the age and condition of the patient, and may be individually adjusted.

#### DIAGNOSTIC REAGENTS

5

10

15

20

25

30

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may be used in a wide variety of ways in addition to gene therapy and protein replacement therapy. They may be useful as diagnostic reagents to measure normal or abnormal activity of BRCA2 at the DNA, RNA, and protein level. The present invention therefore encompasses the diagnostic reagents derived from the BRCA2 cDNA and protein sequences as set forth in SEQ. ID. Nos: 4-13. These reagents may be utilized in methods for monitoring disease progression, for determining patients suited for gene and protein replacement therapy, or for detecting the presence or quantifying the amount of a tumor growth inhibitor following such therapy. Such methods may involve conventional histochemical techniques, such as obtaining a tumor tissue from the patient, preparing an extract and testing this extract for tumor growth or metabolism. For example, the test for tumor growth may involve measuring abnormal BRCA2 activity using conventional diagnostic assays, such as Southern, Northern, and Western blotting, PCR, RT-PCR, and immunoprecipitation. In biopsies of tumor tissues, the loss of BRCA2 expression in tumor tissue may be verified by RT-PCR and Northern blotting at the RNA level. A Southern blot analysis, genomic PCR, or fluorescence in situ hybridization (FISH) may also be performed to examine the mutations of BRCA2 at the DNA level. And, a Western blotting, protein truncation assay, or immunoprecipitation may be utilized to analysis the effect at the protein level.

These diagnostic reagents are typically either covalently or non convalently attached to a detectable label. Such a label includes a radioactive label, a colorimetric enzyme label, a fluorescence label, or an

epitope label. Frequently, a reporter gene downstream of the regulatory sequences is fused with the BRCA2 protein or polypeptide to facilitate the detection and purification of the target species. Commonly used reporter genes in BRCA2 fusion proteins include  $\beta$ -galactosidase and luciferase gene.

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular localization of BRCA2 protein in normal and cancerous cell. For example, yeast two-hybrid system has been used in the study of cellular function of BRCA2 to identify the regulator and effector of BRCA2 tumor suppressing function (Sharan et al., Nature 386:804-810 (1997) and Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988)). In addition, the BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be used in *in vivo* cell based and *in vitro* cell free assays to screen natural products and synthetic compounds which may mimic, regulate or stimulate BRCA2 protein function.

#### ANTISENSE INHIBITION

5

10

15

20

25

30

Antisense suppression of endogenous BRCA2 expression may assess the effect of BRCA2 protein on cell growth inhibition using known method in the art (Crooke, *Annu. Rev. Pharmacol. Toxicol.* 32:329-376 (1992) and Robinson-Benion and Holt, *Methods Enzymol.* 254:363-375 (1995)). Given the cDNA sequence as set forth in SEQ ID. NO: 4, 6, 8, 10, and 12, one of skill in the art can readily obtain anti-sense strand of DNA and RNA sequences to interfere with the production of wild-type BRCA2 protein or the mutated form of BRCA2 protein. Alternatively, antisense oligonucleotide may be designed to target the control sequences of BRCA2 gene to reduce or prevent the expression of the endogenous BRCA2 gene.

**ANTIBODIES** 

The BRCA2 protein, polypeptides, or their functional equivalents may be used as immunogens to prepare polyclonal or monoclonal antibodies

capable of binding the BRCA2 derived antigens in a known manner (Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988). These antibodies may be used for the detection of the BRCA2 protein, polypeptides, or a functional equivalent in an immunoassay, such as ELISA, Western blot, radioimmunoassay, enzyme immunoassay, and immunocytochemistry. Typically, an anti-BRCA2 antibody is in solution or is attached to a solid surface such as a plate, a particle, a bead, or a tube. The antibody is allowed to contact a biological sample or a blot suspected of containing the BRCA2 protein or polypeptide to form a primary immunocomplex. After sufficient incubation period, the primary immunocomplex is washed to remove any non-specifically bound species. The amount of specifically bound BRCA2 protein or polypeptide may be determined using the detection of an attached label or a marker, such as a radioactive, a fluorescent, or an enzymatic label. Alternatively, the detection of BRCA2 derived antigen is allowed by forming a secondary immunocomplex using a second antibody which is attached with a such label or marker. The antibodies may also be used in affinity chromatography for isolating or purifying the BRCA2 protein, polypeptides or their functional equivalents.

20

25

30

15

5

10

#### EXAMPLE 1

# <u>Determination of the Coding Sequence Haplotypes of the BRCA2 Gene</u> <u>From Normal Individuals</u>

Approximately 150 volunteers were screened in order to identify individuals with no cancer history in their immediate family (i.e. first and second degree relatives). Each person was asked to fill out a hereditary cancer prescreening questionnaire (See TABLE I). Five of these were randomly chosen for end-to-end sequencing of their BRCA2 gene. A first degree relative is a parent, sibling, or offspring. A second degree relative is an aunt, uncle, grandparent, grandchild, niece, nephew, or half-sibling.

Genomic DNA was isolated from white blood cells of five normal subjects selected from analysis of their answers to the questions above.

Dideoxy sequence analysis was performed following polymerase chain reaction amplification.

All exons of the BRCA2 gene were subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye was attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer® cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator" purchased through ABI.

#### 1. Polymerase Chain Reaction (PCR) Amplification

5

10

20

25

30

Genomic DNA (100 nanograms) extracted from white blood cells of five normal subjects. Each of the five samples was sequenced end to end. Each sample was amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer, 2.5 microliters reverse primer, and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The primers in TABLE II below were used to carry out amplification of the various sections of the BRCA2 gene samples. The primers were synthesized on an DNA/RNA Synthesizer Model 394<sup>®</sup>.

Thirty-five cycles were performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time was increased to 5 minutes, and during the last cycle in which the extension time was increased to 5 minutes.

PCR products were purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

5

10

15

20

25

30

#### 2. Dideoxy Sequence Analysis

Fluorescent dye was attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat # 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator<sup>®</sup>" purchased through ABI.

#### 3. RESULTS

Based upon the sequencing of the five normal individuals, it was determined that the standard sequence found in both GenBank and BIC were inaccurate. In Genbank, a 10 bp stretch (5'-TTTATTTTAG-3') was mistakenly listed as exonic at the 5' end of exon 5 while it should be intronic which would not be included in the cDNA and resultant protein. In addition, a more detrimental error that has the significant potential to lead to an incorrect diagnosis of breast cancer propensity exists in both Genbank and BIC: a sequence of 16 bp (5'-GTGTTCTCATAAACAG-3') should be at the end of exon 15, but instead is listed at the beginning of exon 16 in the database. The disclosure and listing of GenBank is shown in Figure 1. The correct intron/exon sequence of BRCA2 is presented in Figure 2, wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5 (corrected exon 5 is listed as SEQ. ID. NO: 1) and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon 16 (corrected exons 15 and 16 are listed as SEQ. ID. No: 2 and 3 respectively)

The BIC BRCA2 sequence also contains sequence errors in which a strech of nine nucleotides at positions 5554-5460 is listed as CGTTTGTGT (amino acids: Arg-Leu-Cys). The correct sequence at these positions is GTTTGTGTT (amino acids: Val-Cys-Val). In addition, the BIC BRCA2 nuclotides at positions 2024 (codon 599), 4553 (codon 1442), 4815 (codon 1529), 5841 (codon 1871), and 5972 (codon 1915) are T, T, A, C, and T respectively, wherein the correct nucleotides at these positions are C, C, G, T, and C respectively. Among them, the nuclotide errors at codon 599, 1442, 1915 result in amino acids changes.

5

10

15

20

25

Additional differences in the nucleic acids of the five normal individuals were found in ten polymorphic locations. The changes and their positions are found in TABLE III. The individual haplotypes of each chromosome of BRCA2 are displayed in FIGURE 3. In each case, the initial haplotype reported in Genbank (accession number U43746) was subtracted to determine the new haplotypes OMI 1-5. Thus, the Genbank sequence only represents 50% of the haplotypes found; the five new BRCA2 (omi 1-5) DNA sequences are shown as SEQ. ID. NO: 4, 6, 8, 10, and 12, respectively (See FIGURE 3), and the corresponding polypeptides are listed as SEQ. ID. NO: 5, 7, 9, 11, and 13 respectively. In combination, these seven haplotypes represent a functional allele profile for the BRCA2 gene.

The data show that for each of the samples, all exons of BRCA2 were identical except in the region of ten polymorphisms. Six of these polymorphisms were previously identified (Tartigan *et al.*, *Nature Genetics* 12: 333-337 (1996); Phelan *et al.*, *Nature Genetics* 13: 120-122 (1996); Couch *et al.*, *Nature Genetics* 13: 123-125 (1996); Teng, *et al.*, *Nature Genetics* 13: 241-244 (1996); Schubert *et al* 60: 1031-1040 (1997)), but four were unique to this work. Even though the individual polymorphisms may have been identified, none of these complete haplotypes has been previously determined.

WO 99/09164

5

30

40

50

#### TABLE I

#### Hereditary Cancer Pre-Screening Questionnaire

#### Part A: Answer the following questions about your family

- 1. To your knowledge, has anyone in your family been diagnosed with a very specific hereditary colon disease called Familial Adenomatous Polyposis (FAP)?
- 2. To your knowledge, have you or any aunt had breast cancer diagnosed before the age 35?
  - 3. Have you had Inflammatory Bowel Disease, also called Crohn's Disease or Ulcerative Colitis, for <u>more</u> than 7 years?

#### 15 Part B: Refer to the list of cancers below for your responses only to questions in Part B

	Bladder Cancer	Lung Cancer	Pancreatic Cancer
	Breast Cancer	Gastric Cancer	Prostate Cancer
	Colon Cancer	Malignant Melanoma	Renal Cancer
20	Endometrial Cancer	Ovarian Cancer	Thyroid Cancer

- 4. Have your mother or father, your sisters or brothers or your children had any of the listed cancers?
- Have there been diagnosed in your <u>mother</u>'s brothers or sisters, or your <u>mother</u>'s parents <u>more than one</u> of the cancers in the above list?
  - 6. Have there been diagnosed in your <u>father</u>'s brothers or sisters, or your <u>father</u>'s parents <u>more than one</u> of the cancers in the above list?

#### Part C: Refer to the list of relatives below for responses only to questions in Part C

	You	Your mother
	Your sisters or brothers	Your mother's sisters or brothers (maternal aunts
35	& uncles)	
	Your children	Your mother's parents (maternal grandparents)

- 7. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer?

  Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 8. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

#### Part D: Refer to the list of relatives below for responses only to questions in Part D.

45		
	You	Your father
	Your sisters or brothers	Your father's sisters or brothers (paternal aunts
	and uncles)	·
	Your children	Your father's parents (paternal grandparents)

- 9. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer?

  Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 10. Is there a <u>total of 4 or more</u> of any cancers in the list of relatives above other than "simple" skin cancers?
  - © Copyright 1996, OncorMed, Inc.

#### NOT FURNISHED UPON FILING

### NO PRESENTADO/(A) EN EL MOMENTO DE LA PRESENTACIÓN

NON SOUMIS(E) AU MOMENT DU DEPÔT

#### NOT FURNISHED UPON FILING

# NO PRESENTADO/(A) EN EL MOMENTO DE LA PRESENTACIÓN

NON SOUMIS(E) AU MOMENT DU DEPÔT

# TABLE II BRCA2 PRIMER SEQUENCES

SEQ. ID. Number	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
PCR Product Length	263		364		268		453		248		319	143		338		255		621
Oligo	20	41	24	22	44	22	40	38	38	36	40	22	19	40	40	24	37	42
SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	5-TGA GTT TTA CCT CAG TCA CA-3'	5'-CAG GAA ACA GCT ATG ACC CTG TGA CGT ACT GGG TTT TTA GC-3'	3'-GAT CTT TAA CTG TTC TGG GTC ACA-3'	5'-CCC AGC ATG ACA CAA TTA ATG A-3'	5'-TGT AAA ACG ACG GCC AGT AGA ATG CAA ATT TAT AAT CCA GAG TA-3'	5'-ATC AGA TTC ATC TTT ATA GAA C-3'	5-TGT AAA ACG ACG GCC AGT TGT GTT GGC ATT TTA AAC ATC A-3'	5'-CAG GAA ACA GCT ATG ACC CAG GGC AAA GGT ATA ACG CT-3'	5'-TGT AAA ACG ACG GCC AGT TAA GTG AAA TAA AGA GTG AA-3'	5'-CAG GAA ACA GCT ATG ACC AGA AGT ATT AGA GAT GAC-3'	5'-TGT AAA ACG ACG GCC AGT GCC ATA TCT TAC CAC CTT GTG A-3'	5-TTG CAT TCT AGT GAT AAT ATA C-3'	5'-AAT TGT TAG CAA TTT CAA C-3'	5-TGT AAA ACG ACG GCC AGT TGG ACC TAG GTT GAT TGC AGA T-3'	5-CAG GAA ACA GCT ATG ACC TAA ACT GAG ATC ACG GGT GAC A-3'	5-GAA TAA TAT AAA TTA TAT GGC TTA-3'	5-CAG GAA ACA GCT ATG ACC CCT AGT CTT GCT AGT TCT T-3'	5- TGT AAA ACG ACG GCC AGT ARC TGA AGT GGA ACC AAA TGA TAC-3'
Label	BRCA2-2F	BRCA2-2R/M 13R	BRCA2-3FII	BRCA2-3RII	BRCA2-4F/M 13F	BRCA2-4R-1A	BRCA2-5+6F/M13F	BRCA2-5+6R/M13R	BRCA2-7F/M13F	BRCA2-7R/M13R	BRCA2-8F/M13F	BRCA2-8FIA	BRCA2-8RIA	BRCA2-9F/M13F	BRCA2-9R/M13R	BRCA2-10AF	BRCA2-10AR/M13R	BRCA2-10BF/M13F
Exon	2	7	ဇ	က	4	4	5&6	5&6	7	7	8		8	6	6	10A	10A	10B

# TABLE II BRCA2 PRIMER SEQUENCES

SEQ. ID. Number	32	33	¥	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	20	51
PCR Product		508		304		411		349		344		369		368		366		360		326
Oligo Length	44	40	19	40	37	22	22	21	22	22	23	21	23	20	21	21	21	18	22	21
SEQUENCE (5' TO 3') NOTE:  M13 TAIL INCLUDED  M13 FORWARD = TGT AGA ACG GCC AGT  M13 DEVENCE = CAG GAA ACA GCT ATG ACC	5- CAG GAA ACA GCT ATG ACC ACG TGG CAA AGA ATT CTC TGA AGT AA-3'	5'-TGT AAA ACG ACG GCC AGT CAT CTT GAA TCT CAT ACA G-3'	5'-AGA CAG AGG TAC CTG AAT C-3'	5'- TGT AAA ACG ACG GCC AGT TGG TAC TTT AAT TTT GTC ACT T-3'	5'-CAG GAA ACA GCT ATG ACC TGC AGG CAT GAC AGA GAA T-3'	5'-AAG AAG CAA AAT GTA ATA AGG A-3'	5'-CAT TTA AAG CAC ATA CAT CTT G-3'	5'-TCT AGA GGC AAA GAA TCA TAC-3'	5'-CAA GAT TAT TCC TTT CAT TAG C-3'	5'-AAC CAA AAC ACA AAT CTA AGA G-3'	5'-GTC ATT TTT ATA TGC TGC TTT AC-3'	5'-GGT TTT ATA TGG AGA CAC AGG-3'	5'-GTA TTT ACA ATT TCA ACA CAA GC-3'	5-ATC ACA GTT TTG GAG GTA GC-3'	5-CTG ACT TCC TGA TTC TTC TAA-3'	5'-CTC AGA TGT TAT TTT CCA AGC-3'	5'-CTG TTA AAT AAC CAG AAG CAC-3'	5-AGG TAG ACA GCA AGC-3'	5'-GTA ATA TCA GTT GGC ATT TAT T-3'	5'-TGC AGA GGT ACA TCC AAT AAG-3'
Label	BRCA2-10BR/M13R	BRCA2-10CF/M13F	BRCA2-10CRII	BRCA2-11AF-M13	BRCA2-11AR-M13	BRCA2-11BF	BRCA2-11BR	BRCA2-11CF	BRCA2-11CR	BRCA2-11DF	BRCA2-11DR	BRCA2-11EF	BRCA2-11ER	BRCA2-11FF	BRCA2-11FR	BRCA2-11GF	BRCA2-11GR	BRCA2-11HF	BRCA2-11HR	BRCA2-11IF
Exon	108	10C	10C	11	=	11	=	11	11	11	11	17	=	-	=	11	1	+	-	=

# TABLE II BRCA2 PRIMER SEQUENCES

SEQUENC M13 TORWARD = T	SEQUENC M13 FORWARD = T	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT	Oligo Length	PCR Product	SEQ. ID. Number
1	2	M13 REVERSE = CAG GAA ACA GCT ATG ACC	, 6	Length	52
BRCA2-11IR 5'-GAT CAG TAA	5'-GAT CAG TAA	5-GAT CAG TAA ATA GCA AGT CCG-3	17		30
BRCA2-11JF 5'-TAC TGA AAA T	5'-TAC TGA AAA T	A AAA TGA AGA TAA CAA AT-3'	23	477	53
BRCA2-11JR 5'-ATT TTG TTC TT	5'-ATT TTG TTC TT	5-ATT TTG TTC TTT CTT ATG TCA G-3'	22		\$
BRCA2-11KF-M13 5'-TGT AAA ACG A(	5'-TGT AAA ACG A	5-TGT AAA ACG ACG GCC AGT CTA CTA AAA CGG AGC AA-3'	35	382	22
BRCA2-11KR-M13 5'-CAG GAA ACA G	5'-CAG GAA ACA G	5-CAG GAA ACA GCT ATG ACC GTA TGA AAA CCC AAC AG-3'	35		56
BRCA2-11LF 5'-CAC AAA ATA CI	5'-CAC AAA ATA CI	5-CAC AAA ATA CTG AAA GAA AGT G-3'	22	374	22
BRCA2-11LR 5'-GGC ACC ACA GTC TCA ATA G-3'	5'-GGC ACC ACA G	TC TCA ATA G-3'	19		28
BRCA2-11MF 5'-GCA AAG ACC CTA AAG TAC AG-3'	5'-GCA AAG ACC C	TA AAG TAC AG-3'	50	409	29
BRCA2-11MR 5'-CAT CAA ATA TT	5'-CAT CAA ATA TT	5-CAT CAA ATA TTC CTT CTC TAA G-3'	22		09
BRCA2-11NF-M13 5'-TGT AAA ACG AC	5'-TGT AAA ACG AC	5'-TGT AAA ACG ACG GCC AGT GAA AAT TCA GCC TTA GC-3'	35	306	61
BRCA2-11NR-M13 5'- CAG GAA ACA G	5'- CAG GAA ACA G	5'- CAG GAA ACA GCT ATG ACC ATC AGA ATG GTA GGA AT-3'	35		62
BRCA2-110F 5'-GTA CTA TAG CT	5'-GTA CTA TAG CT	5'-GTA CTA TAG CTG AAA ATG ACA A-3'	22	383	63
BRCA2-110R 5'-ACC ACT GGC TAT CCT AAA TG-3'	5'-ACC ACT GGC T	AT CCT AAA TG-3'	20		64
BRCA2-11PF 5'-TGA AGA TAT TTG CGT TGA GG-3'	5'-TGA AGA TAT TT	G CGT TGA GG-3'	20	355	65
BRCA2-11PR 5'-GTC AGC AAA AAC CTT ATG TG-3'	5'-GTC AGC AAA A	AC CTT ATG TG-3'	20		99
BRCA2-11QF 5'-ACG AAA ATT AT	5'-ACG AAA ATT A	5'-ACG AAA ATT ATG GCA GGT TGT-3'	21	337	29
BRCA2-11QR 5'-CTT GTC TTG C	s'-cTT GTC TTG C	5-CTT GTC TTG CGT TTT GTA ATG-3'	21		89
BRCA2-11RF 5'-GCT TCA TAA G	5'-GCT TCA TAA	5-GCT TCA TAA GTC AGT CTC AT-3'	20	360	69
BRCA2-11RR 5'-TCA AAT TCC 1	5'-TCA AAT TCC 1	5-TCA AAT TCC TCT AAC ACT CC-3'	20		70
BRCA2-11SF-M13 5'-TGT AAA ACG	5'-TGT AAA ACG	5-TGT AAA ACG ACG GCC AGT TAC AGC AAG TGG AAA GC-3'	35	458	7.1

## TABLE II BRCA2 PRIMER SEQUENCES

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
=	BRCA2-11SR-M13	5-CAG GAA ACA GCT ATG ACC AAG TTT CAG TTT TAC CAA T-3'	37		72
=	BRCA2-11TF	5-GTT CTT CAG AAA ATA ATC ACT C-3'	22	344	73
=	BRCA2-11TR	5-TGT AAA AAG AGA ATG TGT GGC-3'	21		74
1	BRCA2-11UF-M13	5'-TGT AAA ACG ACG GCC AGT ACT TTT TCT GAT GTT CCT GTG-3'	39	328	75
=		5'-CAG GAA ACA GCT ATG ACC TAA AAA TAG TGA TTG GCA ACA-3'	39		92
12	BRCA2-12F/M13F	5'-TGT AAA ACG ACG GCC AGT AGT GGT GTT TTA AAG TGG TCA AAA-3'	42	391	77
12	BRCA2-12R/M13R	5'-CAG GAA ACA GCT ATG ACC GGA TCC ACC TGA GGT CAG AAT A-3'	40		78
13	BRCA2/13-2F	5:-TAA CAT TTA AGC ATC CGT TAC-3'	21	310	79
13	BRCA2/13-2R	5-AAA CGA GAC TTT TCT CAT ACT GTA TTA G-3'	28		80
14	BRCA2-14F	5'-ACC ATG TAG CAA ATG AGG GTC T-3'	22	391	81
14	BRCA2-14AR	5'-GCT TTT GTC TGT TTT CCT CCA A-3'	22		82
15	BRCA2-15-2F	5'-CCA GGG GTT GTG CTT TTT AAA-3'	21	284	83
15	BRCA2-15FUT/M13-R 5'-CAG	5'-CAG GAA ACA GCT ATG ACC ACT CTG TCA TAA AAG CCA TC-3'	38		84
16	BRCA2-16AF	5-TTT GGT TTG TTA TAA TTG TTT TTA-3'	24	394	82
16	BRCA2-16AR	5'-CCA ACT TTT TAG TTC GAG AG-3'	20		86
17	BRCA2-17F	5-TTC AGT ATC ATC CTA TGT G-3'	19	282	87
17	BRCA2-17AR	5'-AGA AAC CTT AAC CCA TAC TG-3'	20		88
18	BRCA2-18FUT/M13-	5'-TGT AAA ACG ACG GCC AGT GAA TTC TAG AGT CAC ACT TCC-3'	39	275	89
18	BRCA2-18R/M13R	5-CAG GAA ACA GCT ATG ACC TTT AAC TGA ATC AAT GAC TG-3'	38		06
19	BRCA2-19F/M13F	5'-TGT AAA ACG ACG GCC AGT AAG TGA ATA TTT TTA AGG CAG TT-3'	41	355	91

# TABLE II BRCA2 PRIMER SEQUENCES

		SEQUENCE (5' TO 3') NOTE:	Oligo	PCR	SEQ. ID.
Exon	Label	M13 FORWARD = TGT AAA AGG GCC AGT	Length	Product Length	Number
10	RRCA2-19FUT/M13-R 5'-CAG	GAA AC	39		92
2	BRCA2-20F/M13F		38	296	93
2 6		5-CAG GAA ACA GCT ATG ACC ATG TTA AAT TCA AAG TCT CTA-3'	39		94
23	BRCA2-21F/M13F	5'-TGT AAA ACG ACG GCC AGT GGG TGT TTT ATG CTT GGT TCT-3'	39	304	95
2 6	BRCA2-21R/M13R	5'-CAG GAA ACA GCT ATG ACC CAT TTC AAC ATA TTC CTT CCT G-3'	40		96
200	BRCA2-22F-1A	5-AAC CAC ACC CTT AAG ATG A-3'	19	453	26
22	BRCA2-22R-1A	5-GCA TTA GTA GTG GAT TTT GC-3'	50		86
23	BRCA2-23FII	5-TCA CTT CCA TTG CAT C-3'	16	290	66
3 8	BRCA2-23RII	5-TGC CAA CTG GTA GCT CC-3'	17		100
24	BRCA2-24 2F	5-TAC AGT TAG CAG CGA CAA AA-3'	50	373	101
76	BRCA2-24R/M13R	5-CAG GAA ACA GCT ATG ACC ATT TGC CAA CTG GTA GCT CC-3'	38		102
7 2	BRCA2-25F-7/23		50	427	103
22 25	BRCA2-25R-7/23	5-TAC CAA AAT GTG TGG TGA TG-3'	20		104
2 96	BRCA2/26-2F	5-AAT CAC TGA TAC TGG TTT TG-3'	20	230	105
2 80	BRCA2/26-2R	5-TAT ACT TAC AGG AGC CAC AT-3'	20		106
274	BRCA2-27AF-1A	5-CTG TGT GTA ATA TTT GCG-3'	18	495	107
V 20	BRCA2-27AR/M13R	5'-CAG GAA ACA GCT ATG ACG GCA AGT TCT TCG TCA GCT ATT G-3'	40		108
270	BPCA2-27BE/M13F	3-TGT AAA AC3 ACG GCC AGT GAA TTC TCC TCA GAT GAC TCC A-3'	40	417	109
27B	BRCA2-27BR/M13R	5-CAG GAA ACA GCT ATG ACC TCT TTG CTC ATT GTG CAA CA-3'	38		110
212					

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	-	2	က	4	rc	Frequency
1093/289	<u>A</u> AT → <u>C</u> AT	Asn → His	A/A	AC	A/A	A/A	AC	A = .8 C = .2
1342/372	AAT → CAT	Asn → His	A/C	A/A	AC	AC	AC	A = 0.6 C = 0.4
1593/455	TC <u>A</u> → TC <u>G</u>	Ser → Ser	A/A	A/A	AA	A/A	A/G	A = 0.9 G = 0.1
2457/743	CAĪ→CAC	His → His	1/1	C/T	1/1	1/1	C/T	T = 0.8 C = 0.2
2908/894	<u>G</u> TA → <u>A</u> TA	Val → Ile	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1
3199/991	<u>A</u> AC → <u>G</u> AC	Asn → Asp	A/A	A/G	A/A	A/A	A/G	A = 0.8 G = 0.2

TABLE III NORMAL PANEL TYPING

Frequency	A = 0.8 G = 0.2	T = 0.9 C = 0.1	A = 0.8 G = 0.2	G = 0.9 A = 0.1
ro 	AA	1/1	AA	A/G
4	A/G	1/1	A/G	9/9
က	AVA	1/1	AVA	9/9
7	A/G	Т/Т	NG	9/9
~	A/A	C/T	NA	9/9
Amino Acid Change	Lys → Lys	Val → Val	Ser → Ser	Ala → Thr
Nucleotide Change	AA <u>A</u> → AA <u>G</u>	GT <u>I</u> → GT <u>C</u>	TC <u>A</u> → TC <u>G</u>	GCC → ACC
Position nt/codon	3624/1132	4035/1269	7470/2414	9079/2951

PCT/US98/16905 WO 99/09164

## **EXAMPLE 2**

5

10

20

25

30

## Determination Of A Normal Individual Using BRCA2(OMI 1-5) and The Ten Polymorphisms For Reference

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene; a)
- avoiding misinterpretation of normal polymorphisms found in the b) normal population.

Sequencing was carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, the BRCA2 sequences were used for reference and any polymorphic sites seen in the patient were compared to the nucleic acid sequences listed above for normal codons at each polymorphic site. A normal sample is one which is comparable to the BRCA2(oml 1-5) sequences and contains only minor variations which occur at minor polymorphic sites. The allelic variations which 15 occur at each of the polymorphic sites are paired here for reference.

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- CAT (His) and AAT (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)
- CA<u>T</u> (His) and CA<u>C</u> (His) at position 2457 (codon 743)
- GTA (Val) and ATA (Ile) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GT<u>T</u> (Val) and GT<u>C</u> (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a normal variation for a mutation.

All exons of the BRCA2 gene are subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et

al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

## 10 1. Polymerase Chain Reaction (PCR) Amplification

5

15

20

25

30

The PCR primers used to amplify a patient's sample BRCA2 gene are listed in TABLE II. The primers were synthesized on a DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

## 2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both rorward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup>" purchased through ABI. The BRCA2<sup>(ornl 1-5)</sup> sequences were entered sequentially into the Sequence Navigator software as the standards for comparison. The Sequence Navigator software compares the patient sample sequence to each BRCA2 <sup>(ornl 1-5)</sup> standard, base by base. The Sequence Navigator highlights all differences between the standards (omi 1-5) and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2 (omi 1-5) standards against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known normal polymorphisms for position and base change.

## 3. Results

5

10

The patient's BRCA2 sequence was found to be heterozygous at seven nucleotide positions: 1093 (A/C), 1342 (A/C), 1593 (A/G), 2457 (C/T), 2908 (A/G), 3199 (A/G) and 9079 (A/G). In addition, this changes five amino acids in the polypeptide product: Asn to His at codon 289, Asn to His at codon 372, Val to lle at codon 894, Asn to Asp at codon 991, and Ala to Thr at codon 2951. The question arises whether any or all of these changes have significance to the patient.

Comparison of the patient's results to the BRCA (omi 1-5) haplotypes demonstrates that it matches one of the BRCA2 omi standards (#5), and thus the patient sample is interpreted as carrying a normal gene sequence without causing any elevation in their risk status for breast cancer.

## 25 **EXAMPLE 3**

30

## DETERMINING THE PRESENCE OF A MUTATION IN EXON 11 OF THE BRCA2 GENE USING BRCA2(omi1-5)

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA2 gene. A list of mutations of BRCA2 is publicly available in the Breast Cancer Information Core at http://www.nchgr.nih.gov/dir/lab\_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. et al. Nature Genetics 11:238, (1995).

In this example, a mutation in exon 11 is characterized by amplifying the region of the mutation with a primer set which amplifies the region of the mutation. Sequencing was carried out as in Example 1 using a blood sample from the patient in question. Specifically, exon 11 of the BRCA2 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer® cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

15

20

25

10

5

## 1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA2-11Q-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA2-11Q-R, 10 micromolar solution), and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify segment Q of exon 11 (where the mutation 6174delT is found) are as follows:

BRCA2-11Q-5: 5'- ACG' AAA' ATT' ATG' GCA' GGT' TGT-3'

BRCA2-11Q-R: 5'- CTT' GTC' TTG' CGT' TTT' GTA' ATG-3'

30

The primers are synthesized on an DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the

first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

## 2. <u>Dideoxy Sequence Analysis</u>

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup> purchased through ABI. The BRCA2<sup>(omi 1-5)</sup> sequence is entered into the Sequence Navigator software as the Standard for comparison. The Sequence Navigator software compares the sample sequence to the BRCA2<sup>(omi)</sup> standard, base by base. The Sequence Navigator highlights all differences between the BRCA2<sup>(omi)</sup> normal DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2<sup>(omi 1-5)</sup> standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a sequence variation is compared to known normal polymorphisms for position and base change. The ten frequent polymorphisms which occur in BRCA2 are:

30

25

5

10

15

20

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- CAT (His) and AAT (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)

- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ATA (IIe) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GTT (Val) and GTC (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

## 10 3. Results

5

15

20

Using the above PCR amplification and standard fluorescent sequencing technology, the 6174delT mutation may be found. Mutations are noted by the length of non-matching sequence variation. Such a lengthy mismatch pattern occurs with deletions and insertions. This mutation is named in accordance with the suggested nomenclature for naming mutations, Beaudet, A *et al.*, Human Mutation 2:245-248, (1993). The 6174delT mutation at codon 1982 of the BRCA2 gene lies in segment "Q" of exon 11. The DNA sequence results demonstrate the presence of a one base pair deletion of a T at nucleotide 6174 of the BRCA2<sup>(orni 1-5)</sup> sequences. This mutation interrupts the normal reading frame of the BRCA2 transcript, resulting in the appearance of an in-frame terminator (TAG) at codon position 2003. This mutation is, therefore, predicted to result in a truncated, and most likely, non-functional protein.

## **EXAMPLE 4**

## GENERATION OF MONOCLONAL AND POLYCLONAL ANTIBODIES USING GST-BRCA2 FUSION PROTEIN AS AN IMMUNOGEN

DNA primers are used to amplify a fragment of BRCA2 using PCR technology. The product is then digested with suitable restriction enzymes and fused in frame with the gene encoding glutathione S-transferase (GST) in

Escherichia coli using GST expression vector pGEX (Pharmacia Biotech Inc.) The expression of the fusion protein is induced by the addition of isopropyl-β-thiogalactopyranoside. The bacteria are then lysed and the overexpressed fusion protein is purified with glutathione-sepharose beads. The fusion protein is then verified by SDS/PAGE gel and N-terminus protein sequencing. The purified protein

is used to immunize rabbits according to standard procedures described in Harlow & Lane (1988). Polycolonal antibody is collected from the serum several weeks after and purified using known methods in the art. Monoclonal antibodies against all or fragments of BRCA2 protein, polypeptides, or functional equivalents are obtained using hybridoma technology, see also Harlow & Lane (1988). The BRCA2 protein or polypeptide is coupled to the carrier keyhole limpet hemocyanin in the presence of glutaraldehyde. The conjugated immunogen is mixed with an adjuvant and injected into rabbits. Spleens from antibody-containing rabbits are removed. The B-cells isolated from spleen are fused to myeloma cells using polyethylene glycol (PEG) to promote fusion. The hybrids between the myeloma and B-cells are selected and screened for the production of antibodies to immunogen BRCA2 protein or polypeptide. Positive cells are recloned to generate monoclonal antibodies.

## **EXAMPLE 5**

5

10

15

20

25

30

## **DETECTION OF BRCA2 EXPRESSION IN HUMAN TISSUES AND CELL LINES**

The expression of BRCA2 in human tissues is determined using Northern blot analysis. Human tissues include those from pancreas, testis, prostate, ovary, breast, small intestine, and colon are obtained from Clontech Laboratories, Inc., Palo Alto, CA. The poly(A)+ mRNA Northern blots from different human tissues is hybridized to BRCA2 cDNA probes according to manufacture protocol. The expression level is further conformed by RT-PCR using oligo-d(T) as a primer and other suitable primers.

For Northern Blot analysis of cancer cell lines, the human ovarian cancer cell line SKOV-3 and the human breast cancer cell line MCF-7 are obtained from the American Type Culture Collection. Total RNA is prepared by lysing cell in the presence of guanidinium isocyanate. Poly(A)\* mRNA is isolated using the PolyATract mRNA isolation system from Promega, Madison, WI. The isolated RNA is then electrophoresed under denaturing conditions and transferred to Nylon membrane. The probe used for Northern blot is a fragment of BRCA2 sequence obtained by PCR amplification. The probes are labeled with [ $\alpha$ -32P] dCTP using a random-primed labeling kit (Amersham Life Science, Arlington Heights, IL).

## **EXAMPLE 6**

5

10

15

20

25

30

## **EXPRESSION OF THE BRCA2 PROTEIN**

The whole-cell extracts of BRCA2 transfected cells are subjected to immunoprecipitation and immunoblotting to determine the BRCA2 protein level. The BRCA2 protein or polypeptide is immunoprecipitated using anti-BRCA2 antibodies prepared according to Example 4. Samples are then fractionated using SDS/PAGE gel and transferred to nitrocellulose. Western blot of the BRCA2 protein or polypeptide is performed with the indicated antibodies. Antibody reaction is revealed using enhanced chemiluminescence reagents (Dupont New England Nuclear, Boston, MA).

## **EXAMPLE 7**

## USE OF THE BRCA2(oml1-5) GENE THERAPY

The growth of ovarian or breast cancer may be arrested by increasing the expression of the BRCA2 gene where inadequate expression of that gene is responsible for hereditary ovarian or breast cancer. Gene therapy may be performed on a patient to reduce the size of a tumor. The LXSN vector may be transformed with a BRCA2<sup>(oml1-5)</sup> coding sequence as presented SEQ ID NO:4, 6, 8, 10, or 12 or a fragment thereof.

## Vector

The LXSN vector is transformed with a fragment of the wildtype BRCA2<sup>(omi1-5)</sup> coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, or 12. The LXSN-BRCA2<sup>(omi1-5)</sup> retroviral expression vector is constructed by cloning a *Sal* I linkered BRCA2<sup>(omi1-5)</sup> cDNA or fragments thereof into the *Xho* I site of the vector LXSN. Constructs are confirmed by DNA sequencing. See Holt et al., *Nature Genetics* 12: 298-302 (1996). Retroviral vectors are manufactured from viral producer cells using serum free and phenol-red free conditions and tested for sterility, absence of specific pathogens, and absence of replication-competent retrovirus by standard assays. Retrovirus is stored frozen in aliquots which have been tested.

Patients receive a complete physical exam, blood, and urine tests to determine overall health. They may also have a chest X-ray, electrocardiogram, and appropriate radiologic procedures to assess tumor stage.

Patients with metastatic ovarian cancer are treated with retroviral gene therapy by infusion of recombinant LXSN-BRCA2<sup>(omi1-5)</sup> retroviral vectors into peritoneal sites containing tumor, between 10<sup>9</sup> and 10<sup>10</sup> viral particles per dose. Blood samples are drawn each day and tested for the presence of retroviral vector by sensitive polymerase chain reaction (PCR)-based assays. The fluid which is removed is analyzed to determine:

- 1. The percentage of cancer cells which are taking up the recombinant LXSN-BRCA2<sup>(omi1-5)</sup> retroviral vector combination. Successful transfer of BRCA1 gene into cancer cells has been shown by both RT-PCR analysis and *in situ* hybridization. RT-PCR is performed with by the method of Thompson et al., *Nature Genetics* 9: 444-450 (1995), using primers derived from a BRCA2<sup>(omi1-5)</sup> coding sequence as in SEQ ID NO:4, 6, 8, 10, or 12 or fragments thereof. Cell lysates are prepared and immunoblotting is performed by the method of Jensen *et al.*, *Nature Genetics* 12: 303-308 (1996) and Jensen *et al.*, *Biochemistry* 31: 10887-10892 (1992).
- 2. Presence of programmed cell death using APOTAG® in situ apoptosis detection kit (ONCOR, INC., Gaithersburg, Maryland) and DNA analysis.
- Measurement of BRCA2 gene expression by slide immunofluorescence or Western blot.

Patients with measurable disease are also evaluated for a clinical response to LXSN-BRCA2<sup>(omi1-5)</sup> especially those that do not undergo a palliative intervention immediately after retroviral vector therapy. Fluid cytology, abdominal girth, CT scans of the abdomen, and local symptoms are followed.

25

30

5

10

15

20

For other sites of disease, conventional response criteria are used as follows:

- 1. Complete Response (CR), complete disappearance of all measurable lesions and of all signs and symptoms of disease for at least 4 weeks.
- 2. Partial Response (PR), decrease of at least 50% of the sum of the products of the 2 largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. To be considered a PR, no new lesions should have appeared during this period and none should have increased in size.
- 3. Stable Disease, less than 25% change in tumor volume from previous evaluations.

4. Progressive Disease, greater than 25% increase in tumor measurements from prior evaluations. The number of doses depends upon the response to treatment.

## **EXAMPLE 8**

5

10

15

20

25

## PROTEIN REPLACEMENT THERAPY

Therapeutically elevated level of functional BRCA2 protein may alleviate the absence or reduced endogenous BRCA2 tumor suppressing activity. Breast or ovarian cancer is treated by the administration of a therapeutically effective amount of the BRCA2 protein, a polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Clinically effective delivery method is applied either locally at the site of the tumor or systemically to reach other metastasized locations with known protocols in the art. These protocols may employ the methods of direct injection into a tumor or diffusion using time release capsule. A therapeutically effective dosage is determined by one of skill in the art.

Breast or ovarian cancer may be prevented by the administration of a prophylactically effective amount of the BRCA2 protein, polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Individuals with known risk for breast or ovarian cancer are subjected to protein replacement therapy to prevent tumorigenesis or to decrease the risk of cancer. Elevated risk for breast and ovarian cancer includes factors such as carriers of one or more known BRCA1 and BRCA2 mutations, late child bearing, early onset of menstrual period, late occurrence of menopause, and certain high risk dietary habits. Clinically effective delivery method is used with known protocols in the art, such as administration into peritoneal cavity, or using an implantable time release capsule. A prophylactically effective dosage is determined by one of skill in the art.

## TABLE OF REFERENCES

- 1. Sanger, F., et al., J. Mol. Biol. 42:1617, (1980).
- 30 2. Beaucage, et al., Tetrahedron Letters 22:1859-1862, (1981).
  - 3. Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, (1982).
  - 4. Conner, et al., Proc. Natl. Acad. Sci. U.S.A. 80:278, (1983)
  - 5. Saiki, et.al.., *Bio/Technology* 3:1008-1012, (1985)

- 6. Landgren, et al., Science 241:1007,(1988)
- 7. Landgren, et al., Science 242:229-237, (1988).
- 8. PCR. A Practical Approach, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, (1992).
  - 9. Easton et al., American Journal of Human Genetics 52:678-701, (1993).
  - 10. Patent No. 4,458,066.

5

20

30

- 11. Rowell, S., et al., American Journal of Human Genetics 55:861-865, (1994)
- 12. Miki, Y. et al., Science 266:66-71, (1994).
- 10 13. Wooster, R. et al., Science <u>265</u>:2088-2090, (1994).
  - 14. Wooster, R. et al., Nature <u>378</u>:789-792, (1995).
  - 15. Beaudet, A et al., Human Mutation 2:245-248, (1993).
  - 16. Friend, S. et al. Nature Genetics 11:238. (1995).
  - 17. Teng et al, Nature Genetics 13: 241-244 (1996).
- 15 18. Couch et al, Nature Genetics <u>13</u>: 123-125 (1996).
  - 19. Tartigan et al, Nature Genetics 12: 333-337 (1996).
  - 20. Phelan et al, Nature Genetics 13: 120-122 (1996).
  - 21. Schubert et al, American Journal of Human Genetics 60: 1031-1040 (1996).
  - 22. Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989).
  - 23. Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998).
  - 24. Zhang et al., Cell <u>92</u>:433-436 (1998).
  - 25. Sharan et al., Nature 386:804-810 (1997).
  - 26. Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988).
- 25 27. Crooke, Annu. Rev. Pharmacol. Toxicol. <u>32</u>:329-376 (1992)
  - 28. Robinson-Benion and Holt, Methods Enzymol. 254:363-375 (1995).
  - 29. Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.
  - 30. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993.
    - 31. Holt et al., Nature Genetics 12: 298-302 (1996).
    - 32. Thompson et al., Nature Genetics 9: 444-450 (1995).
    - 33. Jensen et al., Nature Genetics 12: 303-308 (1996)
    - 34. Jensen et al., Biochemistry 31: 10887-10892 (1992).

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

5

## SEQUENCE LISTING

5	(1) GENERAL INFORMATION
10	(i) APPLICANT: Murphy, Patricia White, Marga Rabin, Mark Olson, Sheri Yoshikawa, Matthew Jackson, Geoffrey
15	Eskanderi, Tara Schryer, Brenda Park, Michael
	(ii) TITLE OF THE INVENTION: NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE
20	(iii) NUMBER OF SEQUENCES: 111
25	<ul> <li>(iv) CORRESPONDENCE ADDRESS:</li> <li>(A) ADDRESSEE: Howrey &amp; Simon</li> <li>(B) STREET: 1299 Pennsylvania Avenue N.W.</li> <li>(C) CITY: Washington</li> <li>(D) STATE: DC</li> <li>(E) COUNTRY: USA</li> <li>(F) ZIP: 20004</li> </ul>
30	(v) COMPUTER READABLE FORM:
35	(A) MEDIUM TYPE: Diskette (B) COMPUTER: IBM Compatible (C) OPERATING SYSTEM: DOS (D) SOFTWARE: FastSEQ for Windows Version 2.0 (vi) CURRENT APPLICATION DATA:
40	(A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION: (vii) PRIOR APPLICATION DATA:
45	(A) APPLICATION DATA:  (B) FILING DATE:
50	<pre>(viii) ATTORNEY/AGENT INFORMATION:   (A) NAME: Halluin, Albert P   (B) REGISTRATION NUMBER: 25,227   (C) REFERENCE/DOCKET NUMBER: 5371.31.US02</pre>
55	<ul><li>(ix) TELECOMMUNICATION INFORMATION:</li><li>(A) TELEPHONE: 650-463-8109</li><li>(B) TELEFAX: 650-463-8400</li><li>(C) TELEX:</li></ul>
60	<ul><li>(2) INFORMATION FOR SEQ ID NO:1:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 50 base pairs</li></ul>

(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:  (A) NAME/KEY: exon (B) LOCATION: 150 (D) OTHER INFORMATION: Exon 5  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
10 (B) LOCATION: 150 (D) OTHER INFORMATION: Exon 5	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
15 TCCTGTTGTT CTACAATGTA CACATGTAAC ACCACAAAGA GATAAGTCAG	50
(2) INFORMATION FOR SEQ ID NO:2:	
(i) SEQUENCE CHARACTERISTICS:  20 (A) LENGTH: 182 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
25 (ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:	
(A) NAME/KEY: exon (B) LOCATION: 1182 30 (D) OTHER INFORMATION: Exon 15	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	
ATTTAATTAC AAGTCTTCAG AATGCCAGAG ATATACAGGA TATGCGAATT AAGAAGAAAC  35 AAAGGCAACG CGTCTTTCCA CAGCCAGGCA GTCTGTATCT TGCAAAAACA TCCACTCTGC CTCGAATCTC TCTGAAAGCA GCAGTAGGAG GCCAAGTTCC CTCTGCGTGT TCTCATAAAC AG	60 120 180 182
(2) INFORMATION FOR SEQ ID NO:3:	
<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 188 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li></ul>	
45 (D) TOPOLOGY: linear	
<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
(A) NAME/KEY: exon (B) LOCATION: 1188 (D) OTHER INFORMATION: Exon 16	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
CTGTATACGT ATGGCGTTTC TAAACATTGC ATAAAAATTA ACAGCAAAAA TGCAGAGTCT	60
TTTCAGTTTC ACACTGAAGA TTATTTTGGT AAGGAAAGTT TATGGACTGG AAAAGGAATA CAGTTGGCTG ATGGTGGATG GCTCATACCC TCCAATGATG GAAAGGCTGG AAAAGAAGAA TTTTATAG	120 180 188

		(i)	SEC	UENC	CE CI	HARAC	CTERI	STI	CS:								
			(B) 7	CYPE	: nu	10485 clei	c aci	ld	alis								
5		,	(C) 5 (D) 7	TRAI COPOI	LOGY	NESS:	: sir near	ngle									
			i) Mo k) Fl			TYPE	: cDi	AV									
10			(B)	LOC	ATIO	N: 2	29	.104	queno 82 BRCA		MI1)						
15									SEQ								
20	TOTO	CTGC	GC C	TCGG	GTGT	CTT	TTGC TTTT	GGCG GTCA	GTG	GGTC TACT	GCC CCG	GCCG GCCA	GGAG. AAAA	AA G AG A ATG	ACTG CCT	GCGCC AGGGG CACCT ATT Ile	60 120 180 237
25	GGA Gly	TCC Ser 5	AAA Lys	GAG Glu	AGG Arg	Pro	ACA Thr 10	TTT Phe	TTT Phe	GAA Glu	ATT Ile	TTT Phe 15	AAG Lys	ACA Thr	CGC Arg	TGC Cys	285
30	AAC Asn 20	Lys	Ala	Asp	Leu	Gly 25	Pro	Ile	Ser	Leu	Asn 30	Trp	Pne	GIU	GIU	35	333
	TCT Ser	TCA Ser	GAA Glu	GCT Ala	CCA Pro 40	CCC Pro	TAT Tyr	AAT Asn	TCT Ser	GAA Glu 45	CCT Pro	GCA Ala	GAA Glu	GAA Glu	TCT Ser 50	GAA Glu	381
35	CAT His	AAA Lys	AAC Asn	AAC Asn 55	AAT Asn	TAC Tyr	GAA Glu	CCA Pro	AAC Asn 60	CTA Leu	TTT Phe	AAA Lys	ACT Thr	CCA Pro 65	CAA Gln	AGG Arg	429
40	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
45	CAA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	GAT Asp	525
50	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	TAA Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT His	AAA Lys 115	573
	AGT Ser	CTT Leu	CGC Arg	ACA Thr	GTG Val 120	Lys	ACT Thr	AAA Lys	ATG Met	GAT Asp 125	CAA Gln	GCA Ala	GAT Asp	GAT Asp	GTT Val 130	Sei	621
55	TGT Cys	CCA Pro	CTT Leu	CTA Leu 135	Asn	TCT Ser	TGT	CTT Leu	AGT Ser 140	GAA Glu	AGT Ser	CCT Pro	GTT Val	GTT Val 145	Leu	CAA Gln	669
60	TGT Cys	ACA Thr	CAT His	Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg	, Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	Cys	GGG Gly	AGT Ser	717

5	TTG Leu	TTT Phe 165	CAT His	ACA Thr	CCA Pro	Lув	TTT Phe 170	GTG Val	AAG Lys	GGT Gly	CGT Arg	CAG Gln 175	ACA Thr	CCA Pro	AAA Lys	CAT His	765
	Ile 180	Ser	Glu	Ser	Leu	Gly 185	Ala	GIu	vaı	Asp	190	Asp	ATG Met	361		195	813
10	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
15	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
20	TAA TaA	GTG Val	AAA Lys 230	AGC Ser	TAT Tyr	TTT Phe	TCC Ser	AAT Asn 235	CAT His	GAT Asp	GAA Glu	AGT Ser	CTG Leu 240	AAG Lys	AAA Lys	AAT Asn	957
25	GAT Asp	AGA Arg 245	Phe	ATC Ile	GCT Ala	TCT Ser	GTG Val 250	Thr	GAC Asp	AGT Ser	GAA Glu	AAC Asn 255	ACA Thr	AAT Asn	CAA Gln	AGA Arg	1005
	GAA Glu 260	Ala	GCA Ala	AGT Ser	CAT His	GGA Gly 265	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270	GGG Gly	AAT Asn	TCA Ser	TTT	AAA Lys 275	1053
30	GTA Val	AAT	AGC Ser	TGC Cys	AAA Lys 280	GAC Asp	CAC His	ATT Ile	GGA Gly	AAG Lys 285	TCA Ser	ATG Met	CCA Pro	AAT Asn	GTC Val 290	Deu	1101
35	GAA Glu	GAT Asp	GAA	GTA Val 295	Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	Asp	ACC Thr	TCT	GAA Glu	GAA Glu 305	Map	AGT Ser	1149
40	TTT Phe	TC#	TTA Leu 310	Cys	TTT Phe	TCT	AAA Lys	TGT Cys 315	Arg	ACA Thr	AAA Lys	AAT Asn	CTA Leu 320	GIII	AAA Lys	GTA Val	1197
45	AGA Arç	ACT Thi	r Sei	AAC Lys	ACI Thr	AGG Arg	AAA Lys	Lys	ATT	TTC Phe	CAT His	GAA Glu 335	i Ald	AAC Asn	GCT Ala	GAT Asp	1245
	GA/ Glu 340	з Су	r GA	A AAI 1 Ly:	A TCI	Lys 345	Ası	C CAF	A GTO	AAA L Lys	GAA Glu 350	у гув	A TAC	TC#	TTI Phe	GTA Val 355	1293
50	TC'	r GA r Gl	A GT u Va	G GA	A CCA u Pro 360	Ası	GA LAS	r ACT	r GA:	r CCA p Pro 365	о пет	A GAT	r TCA o Ser	CAA A	T GT2 1 Va. 370	A GCA l Ala	1341
55	CA Hi	T CA s Gl	G AA n Ly	G CC s Pr 37	o Ph	r GAG	G AG	T GG r Gl	A AGʻ y Se: 38	r As	C AA	A ATO	C TCC e Sei	2 AA0 2 Lys 38:	S GI	A GTT u Val	1389
60	GT Va	A CC	G TC O Se 39	r Le	G GC u Al	C TG a Cy	T GA s Gl	A TG u Tr 39	p Se	T CA	A CT	A AC u Th	C CT r Let 40	u se	A GG r Gl	T CTA y Leu	1437

	WOS	וכטוכי	.04														
_	AAT Asn	GGA Gly 405	GCC Ala	CAG Gln	ATG Met	Glu	AAA Lys 410	ATA Ile	CCC Pro	CTA Leu	TTG Leu	CAT His 415	ATT Ile	TCT Ser	TCA Ser	TGT Cys	1485
5	Asp 420	Gln	Asn	ATT Ile	Ser	Glu 425	Lys	Asp	ьeu	Leu	430	1111	GIU	A DII	2,5	435	1533
10	Lys	Lys	Asp	TTT Phe	Leu 440	Thr	Ser	GIu	Asn	445	Leu	PIO	Arg	116	450	501	1581
15	Leu	Pro	Lys	TCA Ser 455	Glu	Lys	Pro	Leu	<b>460</b>	GIU	GIU	THE	vai	465	ADII	270	1629
20	Arg	Asp	ي 470	GAG Glu	Gln	His	Leu	Glu 475	Ser	His	Tnr	Asp	480	116	beu	AIG	1677
٥٢	GTA Val	AAG Lys 485	Gln	GCA Ala	ATA Ile	TCT Ser	GGA Gly 490	ACT Thr	TCT Ser	CCA Pro	GTG Val	GCT Ala 495	TCT Ser	TCA Ser	TTT Phe	CAG Gln	1725
25	GGT Gly 500	Ile	AAA Lys	AAG Lys	TCT Ser	ATA Ile 505	TTC Phe	AGA Arg	ATA Ile	AGA Arg	GAA Glu 510	ser	CCT Pro	AAA Lys	GAG Glu	ACT Thr 515	1773
30	TTC Phe	AAT Asn	GCA Ala	AGT Ser	TTT Phe 520	TCA Ser	GGT Gly	CAT His	ATG Met	ACT Thr 525	GAT Asp	CCA Pro	AAC Asn	TTT Phe	AAA Lys 530	٠, ٥	1821
35	GAA Glu	ACT Thr	GAP	GCC Ala 535	Ser	GAA Glu	AGT Ser	GGA Gly	CTG Leu 540	GAA Glu	ATA Ile	CAT His	ACT Thr	GTT Val 545	TGC Cys	TCA Ser	1869
40	CAG Gln	AAC Lys	GA0 Glu 550	a Asp	TCC Ser	TTA	TGT Cys	CCA Pro 555	Asn	TTA Leu	ATT	GAT Asp	AAT Asn 560	GIY	AGC Ser	TGG	1917
	CCA	A GC0 Ala 56	a Thi	ACC Thr	ACA Thr	Gln	Asr	TCT Ser	Val	Ala	шес	The state of the s	Hall	GCA Ala	GGT Gly	TTA Leu	1965
45	ATA 116 580	e Se	C AC	r TTC r Lev	AAA 1 Lys	AAG Lys 585	Lys	A ACA	AAT Asr	AAC Lys	TTT Phe 590	: 116	TAT	GCT Ala	ATA Ile	CAT His 595	2013
50	GA' Asj	r GA p Gl	A AC u Th	A TC	TAT TY1 600	Lys	GG/ Gly	A AAA / Lys	AAA Lys	A ATA	e Pro	G AAZ o Ly:	A GAC	CAA Glr	A AAA 1 Lys 610	A TCA s Ser	2061
55	GA Gl	A CT u Le	A AT	T AAG e As: 61	n Cy	r TC s Se	A GC	C CAC	TT Pho	e GI	A GC	A AA' a As	r GCT n Ala	TTT: A Pho 62!	5 GT	A GCA u Ala	2109
60	CC Pr	A CT	T AC	r Ph	T GC	A AA' a As:	T GC n Al	T GA' a As 63	p Se	A GG r Gl	T TT y Le	A TT u Le	G CA' u Hi: 64	s se	T TC ~ Se	T GTG r Val	2157
	A.A	A AC	A AC	C TG	T TC	A CA	g aa	T GA	T TC	T GA	A GA	A CC	A AC	T TT	G TC	C TTA	2205

	Lys	Arg 645	Ser	Сув	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr	Leu	Ser	Leu	
5	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
10	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
15	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
20	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
25	CCA Pro 740	Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
30	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
35	ACT Thr	CCT Pro	ACT	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	AAC Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
	GGC Gly	Lys	GAA Glu 790	Ser	TAC Tyr	AAA Lys	ATG Met	TCA Ser 795	GAC Asp	AAG Lys	CTC Leu	AAA Lys	GGT Gly 800	AAC Asn	AAT Asn	TAT Tyr	2637
40	GAA Glu	TCT Ser 805	qaA	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
45	GAT Asp 820	Val	TGI Cys	GCT Ala	TTA Leu	AAT Asn 825	Glu	AAT Asn	TAT	AAA Lys	AAC Asn 830	Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
50	CCT Pro	GAA	A AAA	TAC	ATG Met	Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC	2781
55	AA (aA	CA n Gli	AA A	ACA Thi 855	Asn	CTA Leu	A AGA	GTA Val	ATC 11e 860	Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GIU	ACT	2829
	AC'	T TC	A AT: r Ile 87	e Se	A AAF	A ATA	A ACT	GTC Val	Ası	CCA Pro	ASI	TCT Sei	GAA Glu 880	GI	CTI Lev	TTC Phe	2877
60	TC Se	A GA r As	p As	T GAG	G AAT u Ası	AA 7 aA a	r TT	r GTO	C TTO	C CAA	A GTA	A GCT l Ala	TAA 1 13A s	GA/	A AGO	AAT Asn	2925

885 890 895

		885					890					895					
5	AAT Asn 900	CTT Leu	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
10	TGT Cys	GTA Val	AAC Asn	GAA Glu	CCC Pro 920	ATT Ile	TTC Phe	AAG Lys	AAC Asn	TCT Ser 925	ACC Thr	ATG Met	GTT Val	TTA Leu	TAT Tyr 930	GGA Gly	3021
	GAC Asp	ACA Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT Ile	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
15	GTT Val	TAT	GTT Val 950	CTT Leu	GCA Ala	GAG Glu	GAG Glu	AAC Asn 955	AAA Lys	AAT Asn	AGT Ser	GTA Val	AAG Lys 960	CAG Gln	CAT His	ATA Ile	3117
20	AAA Lys	ATG Met 965	ACT Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	AAT Asn	ATA Ile	3165
25	GAT Asp 980	Lys	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	A Asn	GAT Asp	TAC Tyr	ATG Met 990	AAC Asn	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
30	CTC Lev	TTA	GGT Gly	Pro	ATT Ile 1000	Ser	TAA Asn	CAC His	Ser	TTT Phe 1005	GIY	GGT	AGC Ser	Pne	AGA Arg 1010	ACA Thr	3261
	GCT Ala	TCA Ser	AAT Asn	AAG Lys 1015	Glu	ATC	AAG Lys	Leu	TCT Ser 1020	GIu	CAT His	AAC Asn	ATT	AAG Lys 1025	Lys	AGC Ser	3309
35	AA! Lys	A ATO	TTC Phe	Phe	AAA Lys	GAT Asp	ATT	GAA Glu	Glu	CAA Gln	TAI Tyr	CCT Pro	ACT Thr 1040	Set	TTA Leu	GCT Ala	3357
40	TG' Cy	r GT s Va:	l Glu	A ATI	G.F Val	AAT Asr	ACC Thr	Lev	GCA Ala	TTA Leu	GAT Asp	AAT Asn 1055	GIL	AAG Lys	AAA Lys	CTG Leu	3405
45	AG Se 106	r Ly	G CC	CAC Glr	TC!	ATT : Ile	Asr	r ACI	r GTA Val	Ser	Ala	A CAT a His	Lec	CAG Glr	1 361	Ser 1075	3453
50	GT Va	A GT l Va	T GT	r TC:	C GIA	р Су:	r AA/ s Lys	A AA?	r AGT	CAT His	3 II(	A ACC	C CCI	CAC Glr	ATO Met 1090	TTA Leu	3501
	TT Ph	T TC e Se	C AA	G CAG s Gl	n As	r TT	r AA' e As	T TC	A AA( r As: 110(	n Hi	r aa' s as:	r TT n Le	A AC	A CC' r Pro 110	) se.	C CAA	3549
55	A# L}	G GC 's Al	A GA a Gl 111	u Il	T AC e Th	A GA r Gl	A CT u Le	T TC u Se 111	r Th	T AT	A TT e Le	A GA u Gl	A GA u Gl	u se	A GG r Gl	A AGT y Ser	3597
60	C <i>I</i> G	AG TI in Ph 112	ne Gl	A TT u Ph	T AC	T CA r Gl	G TT n Ph 113	e Ar	A AA g Ly	A CC s Pr	A AG	C TA r Ty 113	L II	A TT e Le	G CA u Gl	G AAG n Lys	3645

5	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1155 1150 1155	3693 3741
	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	
10	TCG ATT GGT CAG GTA GAC AGC AGC CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
15	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
20	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
25	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
30	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
35	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
40	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
45	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1315	4173
	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
50	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
55	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1360	4317
60	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG  Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln  1365 1370 1375	4365

	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
5	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
10	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
15	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
20	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605
0.5	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470	4653
25	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
30	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
35	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
40	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555	4893
45	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
50	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
55	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
60	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG	5133

	Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe 1620 1625 1630 1	Leu 1635
5	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser 1640 1645 1650	CCT 5181 Pro
10	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn 1655 1660 1665	TCA 5229 Ser
15	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser 1670 1675 1680	CAG 5277 Gln
	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe 1685 1690 1695	GAT 5325 Asp
20	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr 1700 1705 1710	TTG 5373 Leu 1715
25	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn 1720 1725 1730	CAT 5421 His
30	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser 1735 1740 1745	AAC 5469 Asn
35	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr 1750 1755 1760	CTC 5517 Leu
	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn 1765 1770 1775	GTT 5565 Val
40	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val 1780 1785 1790	AAA 5613 Lys 1795
45	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val 1800 1805 1810	Glu
50	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala 1815 1820 1825	ATT 5709
55	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro 1830 1835 1840	GCA 5757 Ala
	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glv Thr 1845 1850 1855	ATT 5805
60		AAG 5853

	1860		1865	:	1870	1875	
5	GAA AAC Glu Asn	AAC GAG AA Asn Glu As 188	n Lys Ser	AAA ATT TGC Lys Ile Cys 1885	CAA ACG AAA Gln Thr Lys	ATT ATG GCA Ile Met Ala 1890	5901
10	GGT TGT Gly Cys	TAC GAG GC Tyr Glu Al 1895	A TTG GAT a Leu Asp	GAT TCA GAG Asp Ser Glu 1900	GAT ATT CTT Asp Ile Leu	CAT AAC TCT His Asn Ser .905	5949
3.5	CTA GAT Leu Asp	AAT GAT GA Asn Asp Gl 1910	u Cys Ser	ACG CAT TCA Thr His Ser 1915	CAT AAG GTT His Lys Val 1920	TTT GCT GAC Phe Ala Asp	5997
15	ATT CAG Ile Gln 1925	Ser Glu Gl	A ATT TTA u Ile Leu 1930	CAA CAT AAC Gln His Asn	CAA AAT ATG Gln Asn Met 1935	TCT GGA TTG Ser Gly Leu	6045
20	GAG AAA Glu Lys 1940	GTT TCT AA Val Ser Ly	A ATA TCA s Ile Ser 1945	Pro Cys Asp	GTT AGT TTG Val Ser Leu 1950	GAA ACT TCA Glu Thr Ser 1955	6093
25	GAT ATA	TGT AAA TG Cys Lys Cy 196	s Ser Ile	GGG AAG CTT Gly Lys Leu 1965	CAT AAG TCA His Lys Ser	GTC TCA TCT Val Ser Ser 1970	6141
30	GCA AAT Ala Asn	ACT TGT GG Thr Cys Gl 1975	G ATT TTT y lle Phe	AGC ACA GCA Ser Thr Ala 1980	AGT GGA AAA Ser Gly Lys	TCT GTC CAG Ser Val Gln 1985	6189
	GTA TCA Val Ser	A GAT GCT TO Asp Ala Se 1990	r Leu Gln	AAC GCA AGA Asn Ala Arg 1995	CAA GTG TTT Gln Val Phe 2000	TCT GAA ATA Ser Glu Ile	6237
35	GAA GAT Glu Asp 2005	Ser Thr Ly	G CAA GTC s Gln Val 2010	TTT TCC AAA Phe Ser Lys	GTA TTG TTT Val Leu Phe 2015	AAA AGT AAC Lys Ser Asn	6285
40	GAA CAT Glu His 2020	T TCA GAC CA S Ser Asp Gl	G CTC ACA n Leu Thr 2025	Arg Glu Glu	AAT ACT GCT Asn Thr Ala 2030	ATA CGT ACT Ile Arg Thr 2035	6333
45	CCA GAA Pro Gli	A CAT TTA AT 1 His Leu Il 204	e Ser Gln	AAA GGC TTT Lys Gly Phe 2045	TCA TAT AAT Ser Tyr Asn	GTG GTA AAT Val Val Asn 2050	6381
50	TCA TCT Ser Ser	r GCT TTC TO r Ala Phe Se 2055	T GGA TTT r Gly Phe	AGT ACA GCA Ser Thr Ala 2060	AGT GGA AAG Ser Gly Lys	CAA GTT TCC Gln Val Ser 2065	6429
	ATT TT	A GAA AGT TO u Glu Ser Se 2070	r Leu His	AAA GTT AAG Lys Val Lys 2075	GGA GTG TTA Gly Val Leu 2080	GAG GAA TTT Glu Glu Phe	6477
55	GAT TT. Asp Le	u Ile Arg T	T GAG CAT ir Glu His 2090	Ser Leu His	TAT TCA CCT Tyr Ser Pro 2095	ACG TCT AGA Thr Ser Arg	6525
60	CAA AA Gln As 2100	T GTA TCA A	AA ATA CTI /s Ile Leu 2105	CCT CGT GT	GAT AAG AGA Asp Lys Arg 2110	AAC CCA GAG Asn Pro Glu 2115	6573

5	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
10	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
10	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
15	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
20	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
25	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
30	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
35	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
40	GAG GAA ATG GTT T <sup>G</sup> TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
45	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
50	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
50	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
55	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
60	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2355 2350 2355	7293

5	TTT CTG TCT Phe Leu Ser	C AAA TCT CAT Lys Ser His 2360	Leu Tyr Glu	CAT CTG ACT T His Leu Thr Le 2365	TG GAA AAA TCT eu Glu Lys Ser 2370	7341
3	TCA AGC AAT Ser Ser Asn	TTA GCA GTT Leu Ala Val 2375	TCA GGA CAT Ser Gly His 2380	CCA TTT TAT CA	AA GTT TCT GCT In Val Ser Ala 2385	7389
10	ACA AGA AAT Thr Arg Asr 2390	ı Glu Lys Met	AGA CAC TTG Arg His Leu 2395	ATT ACT ACA GO Ile Thr Thr G	ly Arg Pro Thr	7437
15	Lys Val Phe 2405	e Val Pro Pro	Phe Lys Thr 2410	AAA TCA CAT T Lys Ser His Pl 2415	ne His Arg Val	7485
20	Glu Gln Cys 2420	Val Arg Asn 2425	. Ile Asn Leu	Glu Glu Asn A: 2430	2435	7533
25	Asn Ile Asp	Gly His Gly 2440	Ser Asp Asp	AGT AAA AAT A Ser Lys Asn L 2445	ys Ile Asn Asp 2450	7581
	Asn Glu Ile	e His Gln Phe 2455	Asn Lys Asn 2460	Asn Ser Asn G	2465	7629
30	Val Thr Phe 2470	e Thr Lys Cys )	Glu Glu Glu 2475	CCT TTA GAT T Pro Leu Asp Le	eu Ile Thr Ser 30	7677
35	Leu Gln Asr 2485	n Ala Arg Asp	Ile Gln Asp 2490	Met Arg Ile Ly 2495		7725
40	Arg Gln Arg 2500	y Val Phe Pro 2505	Gln Pro Gly	Ser Leu Tyr Lo 2510	2515	7773
45	Ser Thr Le	ı Pro Arg Ile 2520	e Ser Leu Lys	GCA GCA GTA G Ala Ala Val G 2525	ly Gly Gln Val 2530	7821
	Pro Ser Ala	a Cys Ser His 2535	Lys Gln Leu 2540	Tyr Thr Tyr G	2545	7869
50	His Cys Ile 255	e Lys Ile Asr O	Ser Lys Asn 2555	Ala Glu Ser P 25	60	7917
55	Thr Glu As 2565	p Tyr Phe Gly	/ Lys Glu Ser 2570	Leu Trp Thr G 2575	GA AAA GGA ATA ly Lys Gly Ile	7965
60			Trp Leu Ile		AT GGA AAG GCT sp Gly Lys Ala 2595	8013
					CA GGT GTG GAT	8061

Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala

2840 2845 2850

	2840		2845	2850	
5	AT GTG GAG GCC yr Val Glu Ala 2855				
10	AG GAG GAA TTT ln Glu Glu Phe 70		Glu Glu Asn		
15	CA TCA CGT GCA ro Ser Arg Ala				
	CA GAG CTT TAT la Glu Leu Tyr 2905	Glu Ala Val			la
20	AG GGT TAT TTC lu Gly Tyr Phe 2920	Ser Glu Glu			
25	AA ATG TTG AAT ln Met Leu Asn 2935				
30	AG GCC ATG GAA ys Ala Met Glu 50		Gln Lys Glu		
35	TC ACA ACC GTG al Thr Thr Val				
	AA GAT TCA GTT ys Asp Ser Val 2985	Ile Leu Ser			ap
40	CT CTG TTA ACA er Leu Leu Thr 3000	Glu Gly Lys			
45	CA AAA TCT AAA er Lys Ser Lys 3015				
50	CA AAA AAA ACT hr Lys Lys Thr 30		Gln Leu Pro		
55	TT CAG ATT TAC he Gln Ile Tyr				
	AT CCA GAC TTT sp Pro Asp Phe 3065	Gln Pro Ser			le
60	TC GTT TCT GTT al Val Ser Val 3080	. Val Lys Lys			

5	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
10	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120	9597
10	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
15	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
20	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
25	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
30	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
30	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
35	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
40	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
45	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	10029
50	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
30	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
55	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys 3300 3305 3310 3315	10173
60	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	10221

5	TCT Ser		Leu					Ile					Leu			ATA Ile	10269
J	AAT Asn	Thr					Ser					Glu					10317
10	Ser					Thr					Thr					TAT Tyr	10365
15	CTC Leu 3380				Arg					Ser					Gln		10413
20	AGT Ser			Ala					Cys					Gln		ACA Thr	10461
	ATT Ile		Thr					TAA									10485
25																	
(2) INFORMATION FOR SEQ ID NO:5:  (i) SEQUENCE CHARACTERISTICS:																	
30	(i) SEQUENCE CHARACTERISTICS:  30 (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear																
35 (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal																	
			(i) 5	SROUE	NCE	DESC	RIPT			O I D	NO:5	5 :					
40	Mot	(x	•	SEQUE				CION	SEÇ				Clu.	Tle	Dhe	Luc	
40	Met 1	(x Pro	Ile	Gly	Ser 5	Lys	Glu	TION:	SE(	Thr 10	Phe	Phe			15		
40		(x Pro	Ile	Gly	Ser 5	Lys	Glu	TION:	SE(	Thr 10	Phe	Phe			15		
	1	(x Pro Arg	Ile Cys Leu	Gly Asn 20	Ser 5 Lys	Lys Ala	Glu Asp	FION: Arg Leu Pro	SEC Pro Gly 25	Thr 10 Pro	Phe Ile	Phe Ser	Leu Glu	Asn 30	15 Trp	Phe	
40 45	1 Thr Glu Glu	(x Pro Arg Glu Ser	Ile Cys Leu 35	Gly Asn 20 Ser	Ser 5 Lys Ser	Lys Ala Glu	Glu Asp Ala Asn	Arg Leu Pro	Pro Gly 25 Pro	Thr 10 Pro Tyr	Phe Ile Asn	Phe Ser Ser	Leu Glu 45	Asn 30 Pro	15 Trp Ala	Phe Glu	
	1 Thr Glu Glu	Pro Arg Glu Ser	Ile Cys Leu 35 Glu	Gly Asn 20 Ser	Ser 5 Lys Ser Lys	Lys Ala Glu Asn	Glu Asp Ala Asn 55	Arg Leu Pro 40 Asn	Pro Gly 25 Pro Tyr	Thr 10 Pro Tyr Glu	Phe Ile Asn Pro	Phe Ser Ser Asn 60	Leu Glu 45 Leu	Asn 30 Pro	15 Trp Ala Lys	Phe Glu Thr	
	1 Thr Glu Glu	Pro Arg Glu Ser 50 Gln	Ile Cys Leu 35 Glu Arg	Gly Asn 20 Ser His	Ser 5 Lys Ser Lys	Lys Ala Glu Asn Ser 70	Glu Asp Ala Asn 55 Tyr	Arg Leu Pro 40 Asn	Pro Gly 25 Pro Tyr Gl.	Thr 10 Pro Tyr Glu Leu	Phe Ile Asn Pro Ala 75	Phe Ser Ser Asn 60 Ser	Leu Glu 45 Leu Thr	Asn 30 Pro Phe	15 Trp Ala Lys Ile	Phe Glu Thr Ile 80	
45	Thr Glu Glu Pro 65 Phe	Pro Arg Glu Ser 50 Gln Lys	Ile Cys Leu 35 Glu Arg	Gly Asn 20 Ser His Lys	Ser 5 Lys Ser Lys Pro Gly 85	Lys Ala Glu Asn Ser 70 Leu	Glu Asp Ala Asn 55 Tyr	Arg Leu Pro 40 Asn Asn Leu	Pro Gly 25 Pro Tyr Gl	Thr 10 Pro Tyr Glu Leu 90	Phe Ile Asn Pro Ala 75 Tyr	Phe Ser Ser Asn 60 Ser	Leu Glu 45 Leu Thr	Asn 30 Pro Phe Pro	15 Trp Ala Lys Ile Val 95	Phe Glu Thr Ile 80 Lys	
45	Thr Glu Glu Pro 65 Phe Glu	Pro Arg Glu Ser 50 Gln Lys Leu	Ile Cys Leu 35 Glu Arg Glu	Gly Asn 20 Ser His Lys Gln Lys 100	Ser 5 Lys Ser Lys rro Gly 85 Phe	Lys Ala Glu Asn Ser 70 Leu Lys	Glu Asp Ala Asn 55 Tyr Thr Leu	Arg Leu Pro 40 Asn Asn Leu Asp	Pro Gly 25 Pro Tyr Gl Pro Leu 105	Thr 10 Pro Tyr Glu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr	Phe Ser Ser Asn 60 Ser Gln Asn	Leu Glu 45 Leu Thr Ser Val	Asn 30 Pro Phe Pro Pro	15 Trp Ala Lys Ile Val 95 Asn	Phe Glu Thr Ile 80 Lys Ser	
45	Thr Glu Glu Pro 65 Phe Glu Arg	(x Pro Arg Glu Ser 50 Gln Lys Leu His	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115	Gly Asn 20 Ser His Lys Gln Lys 100 Ser	Ser 5 Lys Ser Lys Pro Gly 85 Phe	Lys Ala Glu Asn Ser 70 Leu Lys Arg	Glu Asp Ala Asn 55 Tyr Thr Leu Thr	Arg Leu Pro 40 Asn Asn Leu Asp Val	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys	Thr 10 Pro Tyr Glu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr Arg Lys	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125	Asn 30 Pro Phe Pro Pro 110 Gln	15 Trp Ala Lys Ile Val 95 Asn	Phe Glu Thr Ile 80 Lys Ser Asp	
<b>4</b> 5	Thr Glu Glu Pro 65 Phe Glu Arg	(x Pro Arg Glu Ser 50 Gln Lys Leu His Val	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115 Ser	Gly Asn 20 Ser His Lys Gln Lys 100 Ser Cys	Ser 5 Lys Ser Lys Fro Gly 85 Phe Leu	Lys Ala Glu Asn Ser 70 Leu Lys Arg Leu	Glu Asp Ala Asn 55 Tyr Thr Leu Thr Leu 135	Arg Leu Pro 40 Asn Asn Leu Asp Val 120 Asn	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys	Thr 10 Pro Tyr Glu Leu 90 Gly Thr	Phe Ile Asn Pro Ala 75 Tyr Arg Lys Leu	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125 Glu	Asn 30 Pro Phe Pro Pro 110 Gln ser	15 Trp Ala Lys Ile Val 95 Asn Ala Pro	Phe Glu Thr Ile 80 Lys Ser Asp	
<b>4</b> 5	Thr Glu Glu Pro 65 Phe Glu Arg Asp	(x Pro Arg Glu Ser 50 Gln Lys Leu His Val	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115 Ser	Gly Asn 20 Ser His Lys Gln Lys 100 Ser Cys	Ser 5 Lys Ser Lys Fro Gly 85 Phe Leu	Lys Ala Glu Asn Ser 70 Leu Lys Arg Leu His	Glu Asp Ala Asn 55 Tyr Thr Leu Thr Leu 135	Arg Leu Pro 40 Asn Asn Leu Asp Val 120 Asn	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys	Thr 10 Pro Tyr Glu Leu 90 Gly Thr	Phe Ile Asn Pro Ala 75 Tyr Arg Lys Leu Arg	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125 Glu	Asn 30 Pro Phe Pro Pro 110 Gln ser	15 Trp Ala Lys Ile Val 95 Asn Ala Pro	Phe Glu Thr Ile 80 Lys Ser Asp	
<b>4</b> 5	Thr Glu Glu Pro 65 Phe Glu Arg	(x Pro Arg Glu Ser 50 Gln Lys Leu His Val 130 Leu	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115 Ser Gln	Gly Asn 20 Ser His Lys Gln Lys 100 Ser Cys	Ser 5 Lys Ser Lys Pro Gly 85 Phe Leu Pro	Lys Ala Glu Asn Ser 70 Leu Lys Arg Leu His	Glu Asp Ala Asn 55 Tyr Thr Leu Thr Leu 135 Val	Arg Leu Pro 40 Asn Asn Leu Asp Val 120 Asn Thr	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys Ser	Thr 10 Pro Tyr Glu Leu 90 Gly Thr Cys	Phe Ile Asn Pro Ala 75 Tyr Arg Lys Leu Arg 155	Phe Ser Ser Asn 60 Ser Gln Asn Met Ser 140 Asp	Leu Glu 45 Leu Thr Ser Val Asp 125 Glu Lys	Asn 30 Pro Phe Pro Pro 110 Gln ser	15 Trp Ala Lys Ile Val 95 Asn Ala Pro	Phe Glu Thr Ile 80 Lys Ser Asp Val Val 160	

	Ser	Trp		180 Ser	Ser	Leu	Ala	Thr 200	185 Pro	Pro	Thr	Leu	Ser 205	190 Ser	Thr	Val
5	Leu	Ile 210	195 Val	Arg	Asn	Glu	Glu 215		Ser	Glu	Thr	Val 220		Pro	His	Asp
	Thr 225		Ala	Asn	Val	Lys 230		Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
10	-	_		-	245					250				Glu	255	
			_	260					265					Ser 270		
			275				_	280					285	Ser		
15		290			_		295					300		Thr		
	305	_				310					315			Lys		320
20		-		_	325		_			330				His Glu	335	
				340					345					350 Leu		
25			355					360		_		_	365	Lys		
23		370				-	375				_	380	_	Leu		
	385					390					395			Leu		400
30		-			405					410				Asp	415	
			-	420					425					430 Leu		
35			435					440					445	Glu		
	Val	450 Asn	Lys	Arg	Asp	Glu	455 Glu	Gln	His	Leu	Glu	460 Ser	His	Thr	Asp	Cys
	465 Ile	Leu	Ala	Val	Lys	470 Gln	Ala	Ile	Ser	Gly	475 Thr	Ser	Pro	Val	Ala	480 Ser
40	Ser	Phe	Gln	Gly	485 Ile	Lys	Lys	Ser		490 Phe	Arg	Ile	Arg	Glu	495 Ser	Pro
	Lys	Glu		500 Phe	Asn	Ala	Ser		505 Ser	Gly	His	Met		510 Asp	Pro	Asn
45	Phe		515 Lys	Glu	Thr	Glu		520 Ser	Glu	Ser	Gly	Leu 540	525 Glu	Ile	His	Thr
	Val 545	530 Cys	Ser	Gln	Lys	Glu 550	535 Asp	Ser	Leu	Cys	Pro 555		Leu	Ile	Asp	Asn 560
50		Ser	Trp	Pro	Ala 565		Thr	Thr	Gln	Asn 570		Val	Ala	Leu	Lys 575	
	Ala	Gly	Leu	Ile 580		Thr	Leu	Lys	Lys 585		Thr	Asn	Lys	Phe 590		Tyr
	Ala	Ile	His 595	Asp	Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	11e 605	Pro	Lys	Asp
55	Gln	Lys 610	Ser	Glu	Leu	Ile	Asn 615	Сув	Ser	Ala	Gln	Phe 620	Glu	Ala	Asn	Ala
	625					630					635			Leu		640
60				-	645					650				Glu	655	
	Leu	Ser	Leu	Thr 660	Ser	Ser	Phe	Gly	Thr 665	Ile	Leu	Arg	Lys	Сув 670	Ser	Arg

			675					680					685			Tyr
5	_	690					695					700		Ile		
	705		_			710					715			Glu		720
					725					730				Leu	735	
10				740					745					750		
			755		_			760					765	Ala		
15		770					775					780		Leu		Gly
	785					790					795					800
			_		805					810				Pro	815	
20	=			820					825					Asn 830		
	Leu	Leu	Pro 835	Pro	GIu	гЛя	ıyr	Met 840	Arg	Val	Ala	ser	845	Ser	Arg	гуя
25		850					855					860		Lys		
	Glu 865	Glu	Thr	Thr	Ser	Ile 870	Ser	Lys	Ile	Thr	Val 875	Asn	Pro	Asp	Ser	Glu 880
		Leu	Phe	Ser	Asp 885		Glu	Asn	Asn	Phe 890		Phe	Gln	Val	Ala 895	Asn
30	Glu	Arg	Asn	Asn 900	Leu	Ala	Leu	Gly	Asn 905	Thr	Lys	Glu	Leu	His 910	Glu	Thr
	_		915	-				920					925	Thr		
35		930					935					940		Ser		
	Lys 945	Asp	Leu	Val	Tyr	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960
				-	965					970				Asp	975	
40				980					985					Met 990		
	_		995					1000	)				1005			
45		1010	)				1015	5				1020	)	His		
	Lys 102	-	Ser	Lys	Met	Phe 1030		Lys	Asp	Ile	Glu 1035		Gln	Tyr	Pro	Thr 104
	Ser	Leu	Ala	СЛа	Val 1045		Ile	Val	Asn	Thr 1050		Ala	Leu	Asp	Asn 1055	
50	Lys	Lys	Leu	Ser 1060		Pro	Gln	Ser	Ile 1065		Thr	Val	Ser	Ala 1070		Leu
	Gln	Ser	Ser 107	Val		Val	Ser	Asp 1080	ayD		Asn	Ser	His 1089	Ile	Thr	Pro
55	Gln		Leu		Ser	Lys	Gln 109	Asp		Asn	Ser	Asn 1100	His	Asn	Leu	Thr
55	Pro 110			Lys	Ala	Glu 1110	Ile	-	Glu	Leu	Ser 1115	Thr		Leu	Glu	Glu 112
			Ser	Gln		Glu		Thr	Gln		Arg		Pro	Ser		Ile
60	Leu	Gln	Lys				Glu	Val				Gln	Met	Thr		
	Lys	Thr	Thr	1140 Ser		Glu	Сув	Arg	1145 Asp		Asp	Leu	His	1150 Val		Met

			1155					1160					1165			
			Pro :				7176					1150				
5	Thr	Val	Glu			1100					1177					
	Asn	Lys	Ser .		1205					1210	,				1210	
10			Phe						コンフラ	١				123	,	
			Gln 1235					1240					144	,		
		125	1235 Thr )				1255	5				120	,			
15	-0		Asp			1270	3				1275	)				120
			Val		128	Ξ.				1290	J				123	,
20			Glu	1200	3				1305	5				121	,	
			Arg 1315	•				1320	)				132	2		
		1 2 2	Ser 0				1231	5				134	U			
25	4241	_	Val			1251	<b>`</b>				135:	•				150
			Asn		126	=				137	0				73/	•
30			Gln	120	Λ				138	5				100	•	
			Ala 139	-				3400	ח				140	2		
		4 4 4	Ala 0				747	5				144	U			
35	440	_	Phe			147	ი				143	5				T 4.4
			Phe Asn		144	5				145	.0				143	S .
40			Asn Met	716	^				146	5				14/	U	
			Met 147 Leu	_				74 H	()				740			
							149	15				150				Thr
45	160	\ <del>=</del>				151	0				TOT					152 Lys
					152	5				153	30				153	5 Gly
50				154	L n				154	15				100	··	Lys
			155	:5				156	60				TD	22		Glu
		1 5	70				15'	75				T24	30			Asn
55	15	0.5				159	90				15:	<del>)</del> 5				160 1 Leu
					16	05				16	10				10.	15 s Ser
60				16	20				16	25				10	<b>3</b> U	r Ala
	1.7	e Pn	16		J va			16	40				16	45		

			_					1650	Thr				1.660				
5		Asn	Se				1670	Phe	Tyr			TO / 2					
5	Val	Ser				Ser	Leu	Leu	Glu		1690	}				1037	
				7	ly o	Gln	Pro		Arg	1/02	)				7,10	,	
10			9 9	eu 1	yr	Glu			Ser 1720	)				1/22	,		
			ı H	is I				172	Gln 5				1/40	,			
		Sei	: A				1750	Tyr	His			1/32	,				
15		Туз				776	Asn	Lys	Leu		177	ט				1,,-	,
					Glu	qaA	Gln		Asn	178	<b>-</b>				1,00	,	
20			1	ys 2	Asp	Ala			Tyr 180	0				TOO	_		
		10	l G	lu (	3lu			181	Ser 5				1820	J			
25	100	Ala	a I				183	Ile O	Ser			183	>				TOA
25	Pro	Pro				104	Ile	Ala	Ser		185	()				100	,
					• ~ ~ /	Lys	Val		Asp	125	~				10,	~	
30			-	ys ·	Glu	Asn			Asn 188	n				100	J		
		7.0	t A	lla	Gly			189	Ala 95				Tan	v			
35		As	n S				101	^	Glu			191	5				エフェ
33	Phe	al Al				102	_		Glu		193	U					_
					204	Lys	Val		: Lys	1 44	<b>–</b>				100	•	
40				. ^	Asp	Ile			1 4 5					1/0	-		Ser
			r	Ser	Ala			191	75				T20	•			Lys
45		r Va	1 (				100	`^				144	. –				Phe 200
13	Se	r Gl				ാവ	15				201	LU					
					202	n				202	25				203		Ala
50					Pro	Glu			204	10				204	<u> </u>		Asn
		2.4	al.	Asn	Ser			20	55				200	, ,			Lys -
55	20	<i>-</i> -					20	70				200	/5				208
33	Gl	u G				20	25				20	90				20.	
					216	As:	n Va			21	05				21	10	Arg
60				211	His	з Су			21	20				21	25		r Lys
	G]	lu P	he	Lys	Let	ı Se	r As	n As	n Le	u As	n Va	1 G1	u Gl	y Gl	y Se	r se	r Glu

		2130	)				2135	5				2140				
	0145	Asn	His			2150	)				7122	1				210
5	Asp	Lys	Gln		2165					2170	ı				21/5	)
			Val	2100	١				2185	5				2190	,	
10			Gly 2195					2200	)				220	)		
_		2211	Val				2219	5				2220	,			
	222	=	Ala			2230	)				223	•				227
15			Ser		2249					2250	)				225	•
			Asn	2260	١				226	5				22/	,	
20			Glu 227	=				2286	0				228	5		
		220	Asn 0				229	5				2300	)			
	220	=	Ser His			231	n				231:	•				232
25			Lys		2321	5				233	0				233	9
				234	Λ				234	5				235	U	Leu
30	_		735	_				236	0				236	5		Gln
		227	^				237	5				238	U			Gly
		_				ってす	a				239	∍				240
35	Arg	Pro	Thr		240	5				241	U				241	_
				212	Λ				242	5				243	v	Arg
40			242	=				244	0				244	<b>-</b>		Lys
		245	- ^				245	55				246	U			Gln
	240					247	'n				247	5				Leu 248 Lys
45					248	5				249	U				247	-
	-			250	0.0				250	)5				251	. U	Leu Glv
50			251	15				252	20				254	35		. Gly - Glv
		25	2 / 2				253	35				254	10			Gly
	25	45				255	50				255	5				256
55					256	55				257	70				25	
				25	R O				25	85				25:	<del>,</del> 0	Asp
60			25	95				26	00				26	U5		r Pro
	Gl	-	1 As:	p Pr	o Ly	s Le	u Il 26	e Se: 15	r Ar	g 110	e Trj	o va. 262	20 20	L ASI	r ur	s Tyr

	Arg Trp				2620					2635					
_	2625 Glu Phe			Arg	Сув	Leu			/n 7 U						
5	Lys Tyr	Arg	Tyr	2645 Asp	Thr	Glu	Ile	Asp 2665	Arg	Ser 2	Arg	Arg	Ser 2670	Ala	Ile
	Lys Lys			Glu	Arg	Asp	Asp 2680	Thr	Ala	Ala :	Lys	Thr 2685	Leu	Val	Leu
10	Cys Val	2675 Ser	Asp	Ile	Ile	ser	Leu	Ser	Ala	Asn	Ile 2700	Ser	Glu	Thr	Ser
	269 Ser Asn	O LVs	Thr	Ser	Ser	2695 Ala	Asp	Thr	Gln	Lys	Val	Ala	Ile	Ile	Glu
	2705 Leu Thr														
15															
	Leu Ala		~ ~ ~ 4 /	^				2,14-	•				2.00		
	Leu His		_				2761	1)				2/0-	,		
20	Glu Ala														
	Pro Ala	a Arg			7791	1				2123					
	2785 Phe Pro	o Leu	Pro	Leu	Ser	Ser	Leu	Phe	Ser 281	Asp	Gly	Gly	Asn	Val 281	Gly
25	Cys Va	l Asp			Ile	Gln	Arg	Ala	Tyr	Pro	Ile	Gln	Trp 2830	Met	Glu
	Lys Th	r Ser	282 Ser	0 Gly	Leu	Tyr	Ile	2825 Phe	Arg	Asn	Glu	Arg 284	Glu	Glu	Glu
30	Lys Gl		_				') <b>X</b> A	<i>(</i> )			Lys	Arg	,		
50	28 Leu Ph					705	E .				200	u			
					207	^				20/2	,				
35	2865 Thr Ly														
	Ala Le														
	Asp Pr		Tyr	Leu											
40	Leu As	n Asr	His												
	29 Gln Le	30 u Glu	ı Ile	arg	Lys	Ala	. Met	: Glu	Ser	Ala	Glu	Gln	Lys	Glu	Gln 296
	2945 Gly Le				795	a				27.7				Val	Ser
45	Tyr Se														
			200	0.0				291	(5				200	•	
	Ser Se	201	n =				300	00				300			
50	Tyr H					3 0.	15				304	٠.٠			
	Ile G	ln Le	u Al	a Ala	a Thi 303	c. Ly:	s Ly	s Thi	c Gli	n Tyr 303	Gl:	ı Glı	ı Leı	Pro	304
	3025 Ser A	sp Gl	u Il	e Lev	ı Phe	e Gl	n Il	е Ту	r Gl	n Pro		g Glu	Pro	Let 30!	ı His
55	Phe S	er Ly	s Ph	304 e Let	45 u As)	o Pr	o As	p Ph	30 e Gl:	n Pro	Se	r Cys	s Sei	Gli	
	Asp L		20	c 0				- 40	b 5				50	, ,	
		2.0	75				30	80				30	0.5		
60	-	000				3.0	195				31	UU			e Lys
	Phe T	rp Il	e As	p Le	u As	n Gl	u As	p Il	e Il	е Гу	s Pr	O HI	a we	. ne	u Ile

	2175 312	
	3105 3110 3115 312 Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu 3135 3135	
_	Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu	
5		
	Gly His Phe Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn	
	The Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His He Leu	
10		
	His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser	
	Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu 3205 3210 3215	
	3205 3210 3213 Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu	
15		
	3220  Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met 3245	
	Thr Ser Lys Ser Cys Lys Glu Lys Glu Ile Asp Asp Gln Lys Asn	
20	1955	
20	Cvs Lvs Lvs Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro	
	3270	
	Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys	
	3/90	
25	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile	
	3300 3305 Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe	
	3370	
	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu	
2.0	7775 3340	
30	Ala Ley lle Asp Thr Glp Ala Ley Ley Ser Gly Ser Thr Gly Glu Lys	
	2250 3333	
	Cln Dhe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser	
	22.5	
35	Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys	
	2000	
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys	
	3395	
	Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile	
40	3410 3415	
	(2) INFORMATION FOR SEQ ID NO:6:	
	(i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 10485 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOG: linear	
50	(ii) MOLECULE TYPE: cDNA	
	(ix) FEATURE:	
	(A) NAME/KEY: Coding Sequence	
	(B) LOCATION: 22910482	
55	(D) OTHER INFORMATION: BRCA2 (OMI2)	
رر		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
		60
	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC	120
60		180
	A CARDETTE CARCECCECE GTTTTTTCTCA GCTTACTCCG GCCAPPERE	37
	CTGGAGCGGA CTTATTTACC AAGCATIGGA GGAATATCGI AGGTALLI III	

Met Pro Ile

5	GGA Gly	TCC Ser 5	AAA Lys	GAG Glu	AGG Arg	CCA Pro	ACA Thr 10	TTT Phe	TTT Phe	GAA Glu	ATT Ile	TTT Phe 15	AAG Lys	ACA Thr	CGC Arg	TGC Cys	285
10	AAC Asn 20	AAA Lys	GCA Ala	GAT Asp	TTA Leu	GGA Gly 25	CCA Pro	ATA Ile	AGT Ser	CIT Leu	AAT Asn 30	TGG Trp	TTT Phe	GAA Glu	GAA Glu	CTT Leu 35	333
15	TCT Ser	TCA Ser	GAA Glu	GCT Ala	CCA Pro 40	CCC Pro	TAT Tyr	AAT Asn	TCT Ser	GAA Glu 45	CCT Pro	GCA Ala	GAA Glu	GAA Glu	TCT Ser 50	GAA Glu	381
	CAT His	AAA Lys	AAC Asn	AAC Asn 55	AAT Asn	TAC Tyr	GAA Glu	CCA Pro	AAC Asn 60	CTA Lev	TTT Phe	AAA Lys	ACT Thr	CCA Pro 65	CAA Gln	AGG Arg	429
20	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
25	CAA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	GAT Asp	525
30	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	AAT Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT	AAA Lys 115	573
35	AGT Ser	CTT Leu	CGC Arg	ACA Thr	GTG Val 120	AAA Lys	ACT Thr	AAA Lys	ATG Met	GAT Asp 125	CAA Gln	GCA Ala	GAT Asp	GAT Asp	GTT Val 130	TCC Ser	621
	TGT	CCA Pro	CTT Leu	CTA Leu 135	Asn	TCT Ser	TGT Cys	CTT Leu	AGT Ser 140	Glu	AGT Ser	CCT Pro	GTT Val	GTT Val 145	CTA Leu	CAA Gln	669
40	TGT Cys	ACA Thr	CAT His	Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	Cys	GGG	AGT Ser	717
45	TTG Leu	TTT Phe 165	His	ACA Thr	CCA Pro	AAG Lys	TTT Phe 170	Val	Lys	GGT	Arg	CAG Gln 175	Thr	CCA Pro	AAA Lys	CAT	765
50	ATT Ile 180	Ser	GAA Glu	AGT Ser	CTA	GGA Gly 185	Ala	GAG	GTG Val	GAI Asp	CCT Pro 190	) Asp	ATG Met	TCI Ser	TGG	TCA Ser 195	813
55	AG1 Se1	TCT Ser	TTI Let	A GCT	ACA Thi	Pro	CCC Pro	ACC Thr	CTI Leu	AGT Ser 205	Ser	ACT Thr	CTC Val	CTC Lev	ATA Ile 210	GTC Val	861
60	AG/ Arg	AA A	r GAA	A GA/ 1 Glv 21	Ala د	A TCT	GAI	A ACT	r GT2 r Val 220	l Phe	CCT Pro	CAT His	GAT B Asp	Thi	Thi	GCT Ala	909
60	AA' Asi	r GT( n Vai	G AA	A AG(	C TA'	r Tr	TCC e Se	C AA'	r CA	r GA' s Asj	r GAA	A AG'	r CTC	AA( Ly	AAI E Lys	TAA A	957

230 235 240

			230				:	235					240				
5	Asp	AGA Arg 245	TTT Phe	ATC Ile	GCT Ala	TCT Ser	GTG Val 250	ACA Thr	GAC Asp	AGT Ser	GAA Glu	AAC Asn 255	ACA Thr	AAT Asn	CAA Gln	AGA Arg	1005
10	GAA Glu 260	GCT Ala	GCA Ala	AGT Ser	CAT His	GGA Gly 265	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270	GGG Gly	AAT Asn	TCA Ser	TTT Phe	AAA Lys 275	1053
	GTA Val	AAT Asn	AGC Ser	TGC Cys	AAA Lys 280	GAC Asp	CAC His	ATT Ile	GGA Gly	AAG Lys 285	TCA Ser	ATG Met	CCA Pro	AAT Asn	GTC Val 290	CTA Leu	1101
15	GAA Glu	GAT Asp	GAA Glu	GTA Val 295	TAT Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	GAT Asp	ACC Thr	TCT Ser	GAA Glu	GAA Glu 305	GAT Asp	AGT Ser	1149
20	TTT Phe	TCA Ser	TTA Leu 310	Cys	TTT Phe	TCT Ser	AAA Lys	TGT Cys 315	AGA Arg	ACA Thr	AAA Lys	AAT Asn	CTA Leu 320	CAA Gln	AAA Lys	GTA Val	1197
25	AGA Arg	ACT Thr 325	Ser	AAG Lys	ACT Thr	AGG Arg	AAA Lys 330	TÀS	ATT Ile	TTC Phe	CAT	GAA Glu 335	GCA Ala	AAC Asn	GCT Ala	GAT Asp	1245
30	GAA Glu 340	Cys	GAA Glu	AAA Lys	TCT Ser	AAA Lys 345	Asn	CAA Gln	GTG Val	AAA Lys	GAA Glu 350	пуз	TAC Tyr	TCA Ser	TTT	GTA Val 355	1293
	TCT Ser	GA/	A GTO	GAA L Glu	CCA Pro 360	Asn	GAT Asp	ACT Thr	GAT Asp	CCA Pro 365	Leu	GAT Asp	TCA Ser	AAT Asn	GTA Val 370	GCA Ala	1341
35	CAT His	CAG	G AAG	G CCC F Pro	Phe	GAG Glu	AGT Ser	GGA Gly	AGT Ser 380	Asp	AAA Lys	A ATC	TCC Ser	AAG Lys 385		GTT Val	1389
40	GTA Val	A CC	G TC' o Se: 39	r Lei	GCC Ala	TGT Cyr	GAA Glu	TGG Trp 395	Ser	CAA Gln	CTA Lev	A ACC	C CTT Leu 400		GG1	CTA Leu	1437
45	AA? Ası	r GG n Gl 40	y Al	C CA(	3 ATC	GA(	3 AAA 1 Lys 410	3 116	CCC Pro	CTA Lev	TTO	G CA	9 110	TCT Ser	TC#	A TGT	1485
50	GAG Asj	p Gl	A AA n As	T AT	T TC	A GA r Gl	u Lys	A GAC	C CT	A TTI	A GA 1 As 43	Блп	A GAG	AA( 1 Asi	C AA	A AGA s Arg 435	1533
	AA Ly	G AA s Ly	AA GA /s As	T TT	T CT e Le 44	u Th	T TC	A GAG	B AA' u As	T TC' n Se 44	r Le	G CC	A CG	r AT	T TC e Se 45	T AGC r Ser 0	1581
55	CT Le	'A Co	CA A/ ro L	AA TC /s Se 45	r Gl	G AA u Ly	G CC	A TT.	A AA u As 46	n GI	G GA u Gl	A AC u Th	A GT ir Va	G GT 1 Va 46		T AAG n Lys	1629
60	AC Ar	BA G g A	sp G	AA GA lu Gl 70	AG CA	G CA	T CT is Le	T GA u Gl 47	u se	T CA	T AC	CA GA	AC TG sp Cy 48	0 11	T CI	T GCA u Ala	1677

5	GTA Val	AAG Lys 485	CAG Gln	GCA Ala	ATA Ile	TCT Ser	GGA Gly 490	ACT Thr	TCT Ser	CCA Pro	GTG Val	GCT Ala 495	TCT Ser	TCA ' Ser	TTT Phe	CAG Gln	1725
	GGT Gly 500	ATC Ile	AAA Lys	AAG Lys	TCT Ser	ATA Ile 505	TTC Phe	AGA Arg	ATA Ile	AGA Arg	GAA Glu 510	TCA Ser	CCT Pro	AAA Lys	GAG Glu	ACT Thr 515	1773
10	TTC Phe	AAT Asn	GCA Ala	AGT Ser	TTT Phe 520	TCA Ser	GGT Gly	CAT His	ATG Met	ACT Thr 525	GAT Asp	CCA Pro	AAC Asn	TTT Phe	AAA Lys 530	AAA Lys	1821
15	GAA Glu	ACT Thr	GAA Glu	GCC Ala 535	TCT Ser	GAA Glu	AGT Ser	GGA Gly	CTG Leu 540	GAA Glu	ATA Ile	CAT His	ACT Thr	GTT Val 545	TGC Cys	TCA Ser	1869
20	CAG Gln	AAG Lys	GAG Glu 550	GAC Asp	TCC Ser	TTA Leu	TGT Cys	CCA Pro 555	AAT Asn	TTA Leu	ATT Ile	GAT Asp	AAT Asn 560	GGA Gly	AGC Ser	TGG Trp	1917
25	CCA Pro	GCC Ala 565	Thr	ACC Thr	ACA Thr	CAG Gln	AAT Asn 570	TCT Ser	GTA Val	GCT Ala	TTG Leu	AAG Lys 575	AAT Asn	GCA Ala	GGT Gly	TTA Leu	1965
	ATA Ile 580	Ser	ACT Thr	TTG Leu	AAA Lys	AAG Lys 585	AAA Lys	ACA Thr	AAT Asn	AAG Lys	TTT Phe 590	116	TAT Tyr	GCT Ala	ATA Ile	CAT His 595	2013
30	GAT Asp	GAA Glu	ACA Thr	TCT Ser	TAT Tyr 600	Lys	GGA Gly	AAA Lys	AAA Lys	ATA Ile 605	Pro	AAA Lys	GAC Asp	CAA Gln	AAA Lys 610	TCA Ser	2061
35	GAA Glu	CTA Lev	A ATT	AAC Asn 615	Cys	TCA Ser	GCC	CAG Gln	TTT Phe 620	GIU	GCA Ala	AAT Asn	GCT Ala	TTT Phe 625	GAA Glu	GCA Ala	2109
40	CCA Pro	CTT	T ACA	r Phe	GCA Ala	TAA .	GCT Ala	GAT Asp 635	ser	GGT Gly	TTA	TTG Leu	CAT His 640	501	TCT Ser	GTG Val	2157
45	AA <i>I</i> Lys	A AG	g Se	C TGT r Cys	TCA Ser	CAC Glr	ı Asr	ı Asp	) Ser	GIU	LGIL	A CCA 1 Pro 655	) TITT	TTG Leu	TCC	TTA Leu	2205
	ACT The	r Se	C TC r Se	T TT: r Phe	r GG( e Gly	ACA Thi	r Ile	r CTO	AGO Arg	AAA J Lys	TG: 670	3 561	AGA Arg	TAA . neA <sub>l</sub>	GA#	ACA Thr 675	2253
50	TG Cy	T TC s Se	T AA r As	T AA' n As:	T AC	r Va	A ATO	C TC	r CAC	G GA' n Asj 68	o re	T GAT u Asj	TAT TYT	AAA Lys	GA/ Glu 690	A GCA 1 Ala 0	2301
55	AA Ly	A TG s Cy	T AA s As	T AA In Ly 69	s Gl	A AA u Ly	A CT s Le	A CA	G TT. n Le 70	u Ph	T AT	T ACC	c cca r Pro	A GAA O Glu 705	1 11	r GAT a Asp	2349
60	TC Se	T CI	G TC eu Se	er Cy	C CT	G CA u Gl	G GA n Gl	A GG u Gl 71	y GI	G TG n Cy	T GA s Gl	A AA u As	T GAT n Asj 72	b Pr	A AA o Ly	A AGC s Ser	2397

	WU	77/07	104														
	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT . Asp	Ile	AAA Lys 730	GAA ( Glu (	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
5	CCA Pro 740	GTA Val	CAA Gln	CAT His	Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
10	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
15	ACT Thr	CCT	ACT Thr	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	AAC Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
20	GGC Gly	AAA Lys	GAA Glu 790	TCA Ser	TAC Tyr	AAA Lys	ATG Met	TCA Ser 795	GAC Asp	AAG Lys	CTC Leu	AAA Lys	GGT Gly 800	AAC Asn	AAT Asn	TAT Tyr	2637
	GAA Glu	TCT Ser 805	Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	Lys Lys	TAA Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
25	GAT Asp 820	Va]	A TGI L Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	vaı	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
30	CCT Pro	GAZ Glu	A AAA	TAC Tyr	ATG Met 840	AGA Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	TCA Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC Phe	2781
35	AA( Asr	C CAI	A AAC n Asr	ACA Thr 855	Asn	CTA	AGA Arg	GTA Val	ATC Ile 860	Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GIU	ACT	2829
40	AC.	r TC. r Se	A ATT	TCA Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	Asn	CCA Pro	GAC Asp	TCT Ser	GAA Glu 880	GIU	CTI Lev	TTC Phe	2877
	TC:	A GA r As 88	p Ası	r GAG n Glu	AAT Asn	AAT Asn	TTT Phe 890	. Val	TTC Phe	CAA Glr	GTA Val	A GCT Ala 895	ASI	GAA Glu	AGG Arg	AAT Asn	2925
45	AA As 90	n Le	T GC	T TTA a Lev	A GGA	AAT Asr 905	Thi	AAG Lys	GAA Glu	CTI Let	CAT His 910	3 GIV	A ACA	GAC Asp	Lev	ACT Thr 915	2973
50	TG Cy	T GI	AA AA 1 As	C GAA	920	) Ile	TTC Phe	AAG Lys	AA S	TC: 1 Sei 92!	r Th	C ATO	GTT Val	TTI L Let	TA' 1 Ty: 93	r GGA r Gly	3021
55	GA As	C AC	CA GG ir Gl	T GA' y As; 93	р Гу	A CA	A GC	A ACC	C CA c Gl: 94	n Va	G TC l Se	A AT	T AAI e Ly:	A AA E Ly 94	B AS	T TTG p Leu	3069
60	GT Vá	TT T	yr Va	TT CT al Le	T GC. u Al	A GA a Gl	G GA u Gl	G AA u As: 95	n Ly	AA A aA a	T AG n Se	T GT r Va	A AA 1 Ly 96	8 GI	G CA n Hi	T ATA s Ile	3117
	A	AA A	TG A	CT CT	'A GG	T CA	A GA	TT T	A AA	A TC	G GA	C AT	C TC	C TI	G AA	ATA T.	3165

	Lys Me		Leu	Gly	Gln	Asp 970	Leu	Lys	Ser	qaA	Ile 975	Ser	Leu	Asn	Ile	
5	GAT AA Asp Ly 980	A ATA s Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	AAC Asn	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
10	CTC TT. Leu Le	A GGT u Gly	Pro	ATT Ile 1000	TCA Ser	AAT Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg 1010	ACA Thr	3261
15	GCT TC. Ala Se	r Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser 1020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
20	AAA AT Lys Me	G TTC t Phe 1030	TTC Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr L040	AGT Ser	TTA Leu	GCT Ala	3357
20	TGT GT Cys Va 104	l Glu	ATT Ile	GTA Val	Asn	ACC Thr L050	TTG Leu	GCA Ala	TTA Leu	qaA	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
25	AGC AA Ser Ly 1060	G CCT s Pro	CAG Gln	Ser	ATT Ile 1065	AAT Asn	ACT Thr	GTA Val	Ser	GCA Ala L070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
30	GTA GT Val Va		Ser					Ser					Gln			3501
35	TTT TC Phe Se	r Lys	CAG Gln 1095	GAT Asp	TTT Phe	AAT Asn	Ser	AAC Asn 1100	CAT His	AAT Asn	TTA Leu	Thr	CCT Pro	AGC Ser	CAA Gln	3549
4.0	AAG GC Lys Al	A GAA a Glu 1110	Ile	ACA Thr	GAA Glu	Leu	TCT Ser 1115	ACT Thr	ATA Ile	TTA Leu	Glu	GAA Glu L120	TCA Ser	GGA Gly	AGT Ser	3597
40	CAG TT Gln Ph 112	e Glu	TTT Phe	ACT Thr	Gln	TTT Phe 1130	AGA Arg	AAR Xaa	CCA Pro	Ser	TAC Tyr 1135	ATA Ile	TTG Leu	CAG Gln	AAG Lys	3645
45	AGT AC Ser Th 1140	A TTT r Phe	GAP. Glu	Val	CCT Pro 1145	GAA Glu	AAC Asn	CAG Gln	Met	ACT Thr 1150	ATC Ile	TTA Leu	AAG Lys	Thr	ACT Thr 1155	3693
50	TCT GA Ser Gl	G GAA u Glu	Сув	AGA Arg 1160	GAT Asp	GCT Ala	GAT Asp	Leu	CAT His 1165	GTC Val	ATA Ile	ATG Met	Asn	GCC Ala 1170	CCA Pro	3741
55	TCG AT Ser Il						Ser					Gly				3789
	ATT AA Ile Ly	A CGG s Arg 1190	Lys	TTT Phe	GCT Ala	Gly	CTG Leu 1195	TTG Leu	AAA Lys	AAT Asn	Asp	TGT Cys 1200	AAC Asn	AAA Lys	AGT Ser	3837
60	GCT TO Ala Se	T GGT	TAT	TTA Leu	ACA Thr	GAT Asp	GAA Glu	AAT Asn	GAA Glu	GTG Val	GGG Gly	TTT Phe	AGG Arg	GGC Gly	TTT Phe	3885

1205 1210 1215

5	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
10	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
15	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
	TCT GTC GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1280	4077
20	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
25	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
30	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
35	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317
40	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG  Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln  1365 1370 1375	4365
45	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
50	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
55	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
60	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605

5	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4653
	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
10	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
15	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
20	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA AFT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
25	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1550 1555	4893
	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
30	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
35	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
40	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
45	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val Wis Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
50	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
55	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
60	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325

	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
5	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
10	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
15	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
20	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
23	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
30	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
35	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	575 <b>7</b>
40	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
43	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
50	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 . 1900 1905	5949
55	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	<b>5997</b> -
60	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA	6093

	Glu Lys 1940	Val Se		Ile Ser .945	Pro Cy	s Asp Val 1950		Glu Thr	Ser 955
5	GAT ATA Asp Ile	TGT AA	AA TGT /s Cys 1960	AGT ATA	GGG AA Gly Ly	G CTT CAT s Leu His 1965	AAG TCA Lys Ser	GTC TCA Val Ser 1970	TCT 6141 Ser
10	GCA AAT Ala Asn	ACT TO Thr Cy 197	s Gly	ATT TTT	AGC AC Ser Th	A GCA AGT r Ala Ser 0	Gly Lys	TCT GTC Ser Val 1985	CAG 6189 Gln
15	GTA TCA Val Ser	GAT GG Asp Al	CT TCA la Ser	TTA CAL	AAC GC Asn Al 1995	A AGA CAA a Arg Glr	GTG TTT Val Phe 2000	TCT GAA Ser Glu	ATA 6237 Ile
20	GAA GAT Glu Asp 2005	Ser Ti	CC AAG nr Lys	CAA GTO Gln Val 2010	. Phe Se	C AAA GTA r Lys Val	TTG TTT Leu Phe 2015	AAA AGT Lys Ser	AAC 6285 Asn
20	GAA CAT Glu His 2020	TCA GA	sp Gln	CTC ACA Leu Thi	AGA GA Arg Gl	A GAA AAT u Glu Asn 2030	Thr Ala	Ile Arg	ACT 6333 Thr 035
25	CCA GAA Pro Glu	CAT T	TA ATA eu Ile 2040	TCC CAM	AAA GG Lys Gl	C TTT TCA y Phe Ser 2045	TAT AAT Tyr Asn	GTG GTA Val Val 2050	AAT 6381 Asn
30	TCA TCT Ser Ser	GCT T Ala Pi 20!	ne Ser	GGA TTT	AGT AC Ser Th 206	A GCA AGT r Ala Ser 0	Gly Lys	CAA GTT Gln Val 2065	TCC 6429 Ser
35	ATT TTA	GAA AG Glu Se 2070	GT TCC er Ser	TTA CAC	AAA GT Lys Va 2075	T AAG GGA l Lys Gly	GTG TTA Val Leu 2080	GAG GAA Glu Glu	TTT 6477 Phe
4.0	GAT TTA Asp Leu 2085	Ile A	GA ACT rg Thr	GAG CAT Glu His 2090	s Ser Le	T CAC TAT u. His Tyr	TCA CCT Ser Pro 2095	ACG TCT Thr Ser	AGA 6525 Arg
40	CAA AAT Gln Asr 2100	GTA TO	er Lys	ATA CT Ile Lev 2105	CCT CG	T GTT GAT g Val Asp 2110	Lys Arg	Asn Pro	GAG 6573 Glu 115
45	CAC TGT His Cys	GTA A	AC TCA sn Ser 2120	GAA ATO	GAA AA Glu Ly	A ACC TGC s Thr Cys 2125	AGT AAA Ser Lys	GAA TTT Glu Phe 2130	AAA 6621 Lys
50	TTA TCA	A AAT A Asn A 21	sn Leu	AAT GT Asn Va	r GAA GG L Glu Gl 214	T GGT TCT y Gly Ser	: Ser Glu	AAT AAT Asn Asn 2145	CAC 6669 His
55	TCT ATT	T AAA G Lys V 2150	TT TCT al Ser	CCA TA	r CTC TC r Leu Se 2155	T CAA TTT er Gln Phe	CAA CAA Gln Gln 2160	GAC AAA Asp Lys	CAA 6717 Gln
60	CAG TTO Gln Lev 216	ı Val L	TA GGA eu Gly	ACC AA Thr Ly 217	s Val Se	CA CTT GTT er Leu Val	r GAG AAC l Glu Asn 2175	ATT CAT Ile His	GTT 6765 Val
60	TTG GG Leu Gl	A AAA G y Lys G	AA CAG lu Gln	GCT TC Ala Se	A CCT AL	AA AAC GTA ys Asn Va	A AAA ATG l Lys Met	GAA ATT Glu Ile	GGT 6813 Gly

	2180	:	2185	2190	2195	
5					AAT ATA GAA GTT Asn Ile Glu Val 2210	6861
10					GAA ACA GAA GCA Glu Thr Glu Ala 2225	6909
15	Val Glu			Glu Asp Asp Glu	CTG ACA GAT TCT Leu Thr Asp Ser 2240	6957
13					TGT CCC GAA AAT Cys Pro Glu Asn	7005
20		Met Val Leu		AGA ATT GGA AAA Arg Ile Gly Lys 2270	AGA AGA GGA GAG Arg Arg Gly Glu 2275	7053
25				TCA ATC AAA AGA Ser Ile Lys Arg 2285		7101
30			Ile Glu Asn	CAA GAA AAA TCC Gln Glu Lys Ser 2300	TTA AAG GCT TCA Leu Lys Ala Ser 2305	7149
35	Lys Ser			AAA GAT CGA AGA Lys Asp Arg Arg		7197
35				TGT GTA CCC TTT Cys Val Pro Phe 2335		7245
40		Gln Glu Ile		AAT TTT ACC GCA Asn Phe Thr Ala 2350		7293
45				GAA CAT CTG ACT Glu His Leu Thr 2365		7341
50			Val Ser Gly	CAT CCA TTT TAT His Pro Phe Tyr 2380		7389
E E	Thr Arg		•	TTG ATT ACT ACA Leu Ile Thr Thr		7437
55				ACT AAA TCA CAT Thr Lys Ser His 2415		7485
60		Cys Val Arg		TTG GAG GAA AAC Leu Glu Glu Asn 2430		7533

5	AAC A Asn I			Gly					qaA					Ile		7581
10	AAT G Asn G		lle					Lys					Gln			7629
	GTA A Val T	hr I					Glu					Asp				7677
15	CTT C Leu G 24					Asp					Arg					7725
20	AGG C Arg G 2500				Phe					Ser					Lys	7773
25	TCC A Ser T			Pro					Lys					Gly		7821
30	CCC T Pro S		Ala					Gln					Gly			7869
30	CAT T His C	ys 1					Ser					Ser				7917
35	ACT G Thr G 25					Gly					Trp					7965
40	CAG T Gln L 2580				Gly					Pro					Lys	8013
45	GGA A			Glu					Leu					Gly		8061
50	CCA A		Leu					Trp					Tyr			8109
	ATA T	rp I					Met					Pro				8157
55	AAT A Asn A 26					Pro					Leu					8205
60	TAT 0 Tyr A 2660				Ile					Arg					Lys	8253

5	ATG (			Asp					Lys					Cys			8301
3	GAC A		Ile					Asn					Ser				8349
10	ACT A	Ser 2	Ser 2710	Ala	Asp	Thr	Gln	Lys 2715	Val	Ala	Ile	Ile 2	Glu 2720	Leu	Thr	qaA	8397
15		725	Tyr	Ala	Val	Lys	Ala 2730	Gln	Leu	Asp	Pro	Pro 2735	Leu	Leu	Ala	Val	8445
20	TTA 1 Leu I 2740	ŗÀs	Asn	Gly	Arg	Leu 2745	Thr	Val	Gly	Gln	Lys 2750	Ile	Ile	Leu	His	Gly 2755	8493
0.5	GCA (			Val					Ala					Glu			8541
25	GAA 1 Glu s		Leu					Ser					Arg				8589
30	TGG 1	ryr					Phe					Arg					8637
35		Leu 805	Ser	Ser	Leu	Phe 2	Ser 2810	Asp	Gly	Gly	Asn	Val 2815	Gly	Cys	Val	Asp	8685
40	GTA A Val 1 2820	Ile	Ile	Gln	Arg 2	Ala 2825	Tyr	Pro	Ile	Gln 2	Trp 2830	Met	Glu	Lys	Thr 2	Ser 2835	8733
45	TCT ( Ser (		Leu	Tyr		Phe	Arg	Asn	Glu	Arg	Glu	Glu	Glu	Lys			8781
•	GCA A		Tyr					Gln					Ala				8829
50	AAA 1 Lys :	Ile					Glu					Asn					8877
55	TAT T					Ala					Gln						8925
60	GAT ( Asp ( 2900				Leu					Lys					Pro		8973
	TAC	CTT	GAG	GGT	TAT	TTC	AGT	GAA	GAG	CAG	TTA	AGA	GCC	TTG	AAT	TAA	9021

	Tyr	Leu	Glu	_	Tyr 2920	Phe	Ser	Glu		Gln 2925	Leu	Arg	Ala		Asn 2930	Asn	
5			Gln					Lys					Ile		TTG Leu		9069
10		Arg					Ser					Glu			TTA Leu		9117
15	Arg					Val					Ile				TCA Ser		9165
20					Ser					Ile					TCA Ser		9213
				Leu					Lys					Tyr	CAT His 3010		9261
25			Ser					Lys					Asn		CAG Gln		9309
30		Ala					Gln					Pro			GAT Asp		9357
35	Ile					Tyr					Pro				AGC Ser		9405
40					Asp					CAa					CTA Leu 3		9453
				Val					Lys					Pro	TTC Phe		9501
45			Ser					Asn					Lys		TGG Trp		9549
50		Leu					Ile					Leu			GCA Ala		9597
55	Asn					Pro					Gly				TTA Leu		9645
60					Ser					Ser					CAC His		9693
-															GAC Asp		9741

	VV 77/07104			202000000000000000000000000000000000000
	3	160	3165	3170
5	CTT TGC AAT GAA C Leu Cys Asn Glu 3	GCA GAA AAC AAG CT Ala Glu Asn Lys Le 318	TT ATG CAT ATA CTG C eu Met His Ile Leu F 30 31	CAT GCA AAT 9789 His Ala Asn 185
10	Asp Pro Lys Trp 3190	Ser Thr Pro Thr Ly 3195	AA GAC TGT ACT TCA C ys Asp Cys Thr Ser C 3200	Sly Pro Tyr
	ACT GCT CAA ATC . Thr Ala Gln Ile 3205	ATT CCT GGT ACA GO Ile Pro Gly Thr Gl 3210	GA AAC AAG CTT CTG A ly Asn Lys Leu Leu N 3215	ATG TCT TCT 9885 Met Ser Ser
15	CCT AAT TGT GAG Pro Asn Cys Glu 3220	ATA TAT TAT CAA AG Ile Tyr Tyr Gln Se 3225	GT CCT TTA TCA CTT 3 er Pro Leu Ser Leu ( 3230	TGT ATG GCC 9933 Cys Met Ala 3235
20	Lys Arg Lys Ser	GTT TCC ACA CCT G Val Ser Thr Pro Va 240	TC TCA GCC CAG ATG A al Ser Ala Gln Met 3 3245	ACT TCA AAG 9981 Thr Ser Lys 3250
25	TCT TGT AAA GGG Ser Cys Lys Gly 3255	GAG AAA GAG ATT GA Glu Lys Glu Ile As 326	AT GAC CAA AAG AAC 3 sp Asp Gln Lys Asn 0 60	rgc AAA AAG 10029 Cys Lys Lys 265
30	AGA AGA GCC TTG Arg Arg Ala Leu 3270	GAT TTC TTG AGT AG Asp Phe Leu Ser An 3275	GA CTG CCT TTA CCT ( rg Leu Pro Leu Pro 1 3280	CCA CCT GTT 10077 Pro Pro Val
25	AGT CCC ATT TGT Ser Pro Ile Cys 3285	ACA TTT GTT TCT CO Thr Phe Val Ser Pr 3290	CG GCT GCA CAG AAG C ro Ala Ala Gln Lys A 3295	GCA TTT CAG 10125 Ala Phe Gln
35	CCA CCA AGG AGT Pro Pro Arg Ser 3300	TGT GGC ACC AAA TA Cys Gly Thr Lys T 3305	AC GAA ACA CCC ATA A yr Glu Thr Pro Ile 1 3310	AAG AAA AAA 10173 Lys Lys Lys 3315
40	Glu Leu Asn Ser	CCT CAG ATG ACT CO Pro Gln Met Thr P: 320	CA TTT AAA AAA TTC 1 ro Phe Lys Lys Phe 1 3325	AAT GAA ATT 10221 Asn Glu Ile 3330
45	TCT CTT TTG GAA Ser Leu Leu Glu 3335	AGT AAT TCA ATA G Ser Asn Ser Ile A 33	CT GAC GAA GAA CTT 0 la Asp Glu Glu Leu 2 40 3	GCA TTG ATA 10269 Ala Leu Ile 345
50	Asn Thr Gln Ala 3350	Leu Leu Ser Gly S 3355	CA ACA GGA GAA AAA e er inr Gly Glu Lys e 3360	Gln Phe Ile
55	TCT GTC AGT GAA Ser Val Ser Glu 3365	TCC ACT AGG ACT G Ser Thr Arg Thr A 3370	CT CCC ACC AGT TCA la Pro Thr Ser Ser 3375	GAA GAT TAT 10365 Glu Asp Tyr
ن	CTC AGA CTG AAA Leu Arg Leu Lys 3380	CGA CGT TGT ACT A Arg Arg Cys Thr T 3385	CA TCT CTG ATC AAA hr Ser Leu Ile Lys 3390	GAA CAG GAG 10413 Glu Gln Glu 3395
60	Ser Ser Gln Ala	AGT ACG GAA GAA T Ser Thr Glu Glu C 3400	GT GAG AAA AAT AAG ys Glu Lys Asn Lys 3405	CAG GAC ACA 10461 Gln Asp Thr 3410

ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 

## (2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala

				7	2 4 0					345		Gln			330		
_	Ser	Phe		1 5	Ser (	Glu	Val	Glu	Pro 360	Asn	Asp	Thr	Asp	Pro 365	Leu	Asp	Ser
5		270		a F				375	Phe			Gly	380				
		Glu	. Va				200					Trp 395					
10						4 A E	Ala				4 I U	Ile				110	
					Asp	Gln				425		Asp			430		
15			4 '	rg 1	Lys				440			Glu		445			
10		4 = 4						455				Leu	400				
	465						470					Glu 475					100
20						495					490	Thr				493	
					EAA					505		Arg			210		
25				1 6					520			His		323			
		E 3	^					535				Gly	540				
	- A F						550					Pro 555					
30						565					570	Ser				515	
					580					585		Thr Lys			330		
35			-	0.5					600			Gln		605			
		~ 1	^					615				Asp	020				
4.0	C 2 1	-					- 630	1				635 Asp					
40						645					650	)				033	Arg
					660					665	5				6/0		Tyr
45									680	ì				002	•		Pro
								69	•				700	,			Asp
50	70	_					71(	)				/ 12	,			. Ala	Ala
50						725					73	0			: Asp	Thr	Asp
					740	`				74	5			ı Ası	n Ala	,	Thr
55				755					760	0			ı Se	r Asi	•		L Met
		7	70				s Gl	77 u Se	5			t Se	r Asj	U			s Gly 800
60	70					ı Se	79 r As	0			u Th	r Ly	5				t Glu
	Ľ۷	/s A	sn	Glr	n Asj	80 p Va	s 1 Cy	s Al	a Le	u As	81 n Gl	.u As	n Ty	r Ly	s As		l Glu

									825					830		
	Leu	Leu	Pro	820 Pro	Glu	Lys	Tyr	Met	Arg	Val	Ala	Ser			Arg	Lys
5	Val	Gln	835 Phe	Asn	Gln	Asn	Thr	840 Asn	Leu	Arg	Val	Ile 860	Gln	Lys	Asn	Gln
					Ser	070					0/3	Asn				
					Asp 885	Asn				กรบ						
10				~~~	Leu				405	Thr				210		
				Сув	Va1			4711					127			
15			Gly		Thr		0.2 E					240				
		Asp	Leu		Tyr	$\alpha = \alpha$					7777					
20	Gln	His			Met 965					976						
				~~~	ГЛЗ				985					224		
					Leu			100	11				100.	_		
25			_		Ser Met		יחח	_				102	_			
		_			Met Val	102	Λ				TOO	_				
30					104 Lys	=				בטו	U				~~~	_
					. ~				1 () 6	~				,	•	Pro
				_				108	(1				100	~		
35																Thr
							^				111					Glu 112
40		c Gly				Glu	ı Phe									
40					Thi	Phe			114	+ つ					•	Leu
				s Se	r Glu											Met
45	As		a Pro	se:	r Ile	e Gly	/ Gli	n Val	L As	Ser	c Sei	: Lys 118	Glr 10	Pne	GIU	Gly
			l Gl			11	חם	s Phe			11:	ı Lev 95	Lys			20 Cys
50	As				12	r Gl	у Ту			1.2	Lυ					
50					r Se	r Al			- 12	25					, ,	r Glu
								12	Δ ( )				14.	x -J		e Ser
55							17	55				14				r Lys
						17	חדי				14	, ,				n Asp 128 n Asn
60						OF				1.2	. 90					n Asn 95 u Asn
	As	sn I.	re G1		300	řī Tř	T G1	.,	13	305	· =			13	10	

			Arg 1315	-				7 2 7 11					1323			
5			Ser	His								Asp 1340				
5		Thr	Val			1251	Lys	Asp			1353	Leu				100
	Gln	His			1265	Leu	Lys			137	,	Phe			10/-	,
10				1386	Lys	Glu			138	•		Thr		1320	,	
			120	<b>E</b>				1400	)			Ser	1405	•		
15		943	^				7479	5				Asp 1420	3			
		_				7/2/	n				143	Ile 5				TAA
					7 4 4 1	5				145	U	Gln			143.	,
20				116	Λ				146	5		Ser		T-4 / (	,	
			7 4 7	_				14 X I	L)			Asp	T.4 O.	,		
25		140	^				1491	5				Gly 150	U			
	1	_				151	n				TOT	Ile 5				102
					152	5				153	0	Val			133	,
30				154	٥				154	5		Glu		TOO	U	
			3.55	E				156	O			Ala	720	_		
35							157	Ε.				158 Asn	U			Glu Asn
	2 5 0	-				159	Λ				723	5 Pro				100
4.0					760					161	. U				101	5 Ser
40				162	Λ.				162	. 5				100	Ÿ	Ala
			163	1 5				164	0				164	Þ		Ile
45		16	ΕΛ.				165	55				100	Ü			Ser
	766	==				167	70				167	/5				168 Gly
Ε0					166	15				16:	90				103	. Gly
50				171	n n				170	)5				7/1	. U	qaA ı
			17	1 6				172	20				1/4	25		Ser
55		17	30				173	35				174	Ł U			Ser
	3 77	A E				17	50				17	55				176 L Leu
60					17	<b>65</b>				17	70				Τ,	75 e Ser
60				17	RΛ				17	85				1/2	90	o Ile
	T D		<i>1</i>					-								

			1795					1800					1805			_
			Glu	Glu	Leu		1016					1020				
5	Ala	Ala	Ile		Leu	1020					T073					
	Pro	Pro			Arg 1845					1000	,					
10					) Fàa 1942				INA	1				70,0		
				•	Asn			3 X X ()					100-	,		
					Cys		1 2 9 5					TOOL	,			
15		_			Asp	1010	1				121:	3				
					Gln 1925	-				1 7 5 1	J					•
20				304	Lys 0				1 44.	3				100	,	
			300	-	Ile			7960	}				150.	_		
			_		Asn		107	^				T20	•			
25		_			Ser	1996	n				199	<b>5</b>				200
					Asp 200	_				201	U				201.	,
30				202	His O				202	5				200	•	
			202	_	Glu			204	()				204	J		
		~~~	^		Ser		205	5				206	U			
35		_			Leu	207	^				201	_				200
					Leu 208	=				209	U				2.00	_
40				010	. ^				210	<b>1</b> 5				211	U	Arg
				_				<b>ว</b> וס	n				212			Lys
		211	١.				213	5				2 T-4	: •			Glu
45	22.4					215	S O				213					216
					216	55				217	/0				21,	
50				211	R A				218	35				211	, 0	Met
			21	90				220	00				221	,,		Asn
		22	٦.				22:	15				22	20			Glu
55		<b>2</b> E				22	30				22	35				224
					22	45				22	50				22.	
60				22	60				22	65				22	, 0	s Arg
	Ar	g Gl		u Pr 75	o Le	u Il	e Le	u Va 22	1 Gl 80	y Gl	u Pr	o Se	r II 22	е шу 85	5 Ar	g Asn

		000	Asn				2295					2300	)			
	Lys	Ala	Ser	Lys	Ser	Thr	Pro	Asp	Gly	Thr	Ile	Lys	qaA	Arg	Arg	Leu 232
5	2306	:				2310	)				2315	•				232
			His		2325	:				2330	)				2335	)
			Lys	2340	)				2345	;				2350	)	
10	Gly	Gln	Glu	Phe	Leu	Ser	Lys	Ser 2360	His	Leu	Tyr	Glu	His 2365	Leu	Thr	Leu
	Glu	Lve	2355 Ser	Ser	Ser	Asn	Leu	Ala	Val	Ser	Gly	His			Tyr	Gln
		2370	n				2375	i				2380	)			
			Ala	Thr	Arg			Lys	Met	arg	H18	Leu:	116	1111	IIII	240
15	238	5	Thr	τ	1703	2390	) 1751	Pro	Dro	Phe			Lvs	Ser	His	
					2409	5				241	)				241:	•
	His	Arg	Val	Glu	Gln	ayɔ	Val	Arg	Asn	Ile	Asn	Leu	Glu	Glu	Asn	Arg
				2420	)				2425	5				2430	j	
20			Gln 243					2440	)				244	5		
	Ile	Asn	Asp	Asn	Glu	Ile	His	Gln	Phe	Asn	Lys	Asn	Asn	Ser	Asn	Gln
		245	<b>n</b>				2455	;				2460	J			
	Ala	Ala	Ala	Val	Thr	Phe	Thr	Lys	CAR	Glu	Glu	Glu -	Pro	Leu	Asp	Leu 248
25	246	5			<b>-</b>	2470	)	_		<b>-1</b> -	247		Mot	7 20	Tla	
	Ile	Thr	Ser	Leu			Ala	Arg	Asp	11e	ν GTU	Asp	Mec	Arg	249	шув 5
	_		Gln	•	2489	7 ~~	17-1	Dhe	Dro			Glv	Ser	Leu		
				2500	)				250	5				251	J	
30	Ala	Lys	Thr 251	Ser	Thr	Leu	Pro	Arg 2520	Ile	Ser	Leu	Lys	Ala 252	Ala 5	Val	Gly
	Glv	Gln	Val	Pro	Ser	Ala	Cvs	Ser	His	Lys	Gln	Leu			Tyr	Gly
		253	n				2535	5				254	3			
	Val	Ser	Lys	His	Cys			Ile	Asn	Ser	Lys	Asn	Ala	GIU	ser	256
35	254	5				255	0	Db -	<b>~</b> 3	T 1/0	255		Leu	Trn	Thr	
			His		256	5				257	0				251	2
			Ile	258	0				258	5				259	υ	
40	Gly	Lys	Ala	Gly	Lys	Glu	Glu	Phe	Tyr	Arg	Ala	Leu	Cys	_Asp	Thr	Pro
			259	5				260	0				260	5		
	Gly		Asp	Pro	Lys	Leu	11e	ser -	Arg	116	Trp	262	V TÀT	Wali	nrs	-7-
		261	.0 lle	~ 7	<b></b>	7	261	) 	71-	Mot	Glu			Phe	Pro	Lvs
4.5			lle								263	5 5				264
45	262	:5 Dhe	Ala	Δgn	Ara	Cvs	Leu	Ser	Pro	Glu			Leu	Leu	Gln	Leu
					264	5				265	0				265	כ
	Lys	туг	Arg	Tyr 266		Thr	Glu	Ile	Asp 266	Arg	Ser	Arg	Arg	Ser 267	Ala O	Ile
50	Lvs	LVE	: Ile	Met	Ğlu	Arg	Asp	Asp			Ala	Lys	Thr	Leu	Val	Leu
•			267	5				268	0				268	5		
	Суя		l Ser	Asp	Ile	Ile	Ser 269		Ser	Ala	Asn	11e 270	Ser	Glu	Thr	ser
	0	269	eo n Lys	. The		CAY	ZOJ Ala	a Den	Thr	· Glr	Lvs			ı Ile	Ile	Glu
55	270		п гре	1111	261	271	0	. Aug			271	.5				272
55	Lei	.i Thi	r Asp	Glv	Trr	Tyr	Ala	Val	Lys	Ala			Asp	Pro	Pro	Leu
					272	25				273	30				2/3	5
	Le	u Ala	a Val	Leu 274		a Asr	Gly	Arg	Lev 274	i Thi	. Val	Gly	Glr.	1 Lys 275	: Ile :0	Ile
60	J.e.	и Ніз	s Gly	/ Ala	Glı	ı Let	ı Val	Gly			) Asy	Ala	Cys	s Thr	Pro	Leu
			275	55				276	50				27€	55		
	G1	u Al	a Pro	o Glu	ı Sei	r Lei	ı Met	: Leu	Lys	a Ile	e Sei	r Ala	a Asr	n Ser	Tur	Arg

	2770		2775		2780	
	Pro Ala Arg 2 2785	279	0	279	5	280
5	Phe Pro Leu l	Pro Leu Ser 2805	Ser Leu	Phe Ser Asp 2810	Gly Gly F	Asn Val Gly 2815
	Cys Val Asp V	Val Ile Ile 2820		Ala Tyr Pro 2825		rrp Met Glu 2830
10	Lys Thr Ser S	Ser Gly Leu	Tyr Ile 2840		Glu Arg 0 2845	3lu Glu Glu
	Lys Glu Ala A 2850	Ala Lys Tyr	Val Glu 2 2855	Ala Gln Gln	Lys Arg I 2860	Leu Glu Ala
	Leu Phe Thr 1	Lys Ile Gln 287		Phe Glu Glu 287		Glu Asn Thr 288
15	Thr Lys Pro			Ala Leu Thr 2890	Arg Gln G	Gln Val Arg 2895
	Ala Leu Gln			Tyr Glu Ala 2905		Asn Ala Ala 2910
20	Asp Pro Ala 3		Gly Tyr		Glu Gln I 2925	Leu Arg Ala
20	Leu Asn Asn E	His Arg Gln				Ala Gln Ile
	Gln Leu Glu	lle Arg Lys			Glu Gln I	
25	2945 Gly Leu Ser A	295) Arg Asp Val			•	
	Tyr Ser Lys I	2965	Non Cor 1	2970	Cor Tle T	2975
	-	2980	:	2985	2	2990
30	Ser Ser Asp 1	deu Tyr Ser	3000	imi Giu Giy	3005	.yr Arg rre
	Tyr His Leu A		3015		3020	
	Ile Gln Leu I 3025	3030	כ	303	5	304
35	Ser Asp Glu	3045		3050		3055
		3060		3065	3	3070
40	Asp Leu Ile (	_	3080		3085	
	Pro Phe Val 7		3095		3100	
	Phe Trp Ile A	3110	כ	311	5	312
45	Ala Ala Ser A		Trp Arg	Pro Glu Ser 3130	Lys Ser G	3135
	Thr Leu Phe	3125 Ala Gly Asp 3140				
50	Gly His Phe G			Lys Met Lys		
30	Ile Asp Ile 1	Leu Cys Asn	_			His Ile Leu
	His Ala Asn A	Asp Pro Lys 319	Trp Ser	Thr Pro Thr 319	Lys Asp C	Cys Thr Ser
55	Gly Pro Tyr			Pro Gly Thr 3210	Gly Asn I	Lys Leu Leu 3215
	Met Ser Ser	Pro Asn Cys 3220		Tyr Tyr Gln 3225		Leu Ser Leu 3230
60	Cys Met Ala 1 3235	Lys Arg Lys	Ser Val		Val Ser A	Ala Gln Met
	Thr Ser Lys 3 3250	Ser Cys Lys	Gly Glu 3255	Lys Glu Ile	Asp Asp G	In Lys Asn

	Cys I 3265	_	Lys	Arg	Arg	Ala 3270		Asp	Phe	Leu	Ser 3275	Arg	Leu	Pro	Leu	Pro 328	
_	Pro 1	Pro '	Val	Ser			Сув	Thr	Phe	Val 3290		Pro	Ala	Ala	Gln 3295	Lys	
5	Ala 1	Phe	Gln	Pro	3285 Pro		Ser	Cys	Gly			Tyr	Glu	Thr			
	Lys 1			3300	)				3305	;				3310	)		
- ^	Asn (	_	3315					3320	1				3325	,			
10		3330					3335	;				3340	)				
	Ala 1					3350	)				3355	;				336	
15	Gln 1	Phe	Ile	Ser	Val 3365		Glu	Ser	Thr	Arg 3370		Ala	Pro	Thr	Ser 3375	Ser	
12	Glu i	Asp	Tyr		Arg		Lys	Arg		Сув		Thr	Ser	Leu 3390	Ile		
	Glu	Gln	Glu	3380 Ser	Ser	Gln	Ala	Ser	3385 Thr		Glu	Cys		Lys		Lys	
20	Gln A		3395 Thr		Thr	Thr	Lvs	3400 Lvs		Ile			3405	;			
20		3410					3415		-3-								
			(2)	INE	FORM	(OIT	FOF	SEÇ	) ID	№: ОИ	3 :						
25		(i	) SE	QUE	ICE (	CHAR	CTE	RISTI	CS:								
						1048 iclei			aire	3							
			(C)	STRA	ANDEI	ONESS	3: do	ouble	<b>:</b>								
30						Y: 13											
			i) M x) F			TYPI	E: cI	AAC									
		•-	·			EY: (	rodi.	.a Ce	mier	·ce							
35			(B)	LO	CATIO	ON: 2	229.	104	82								
			(D)	OTI	HER :	INFO	TAMS:	ON:	BRCA	12 (0	OMI3)						
		(x	i) S	EQUI	ENCE	DES	CRIP	CION:	SEÇ	DID	NO:8	3:					
40	GGTG	GCGC	GA C	CTT	CTGA	AA C	FAGG	CGGCI	GAC	GCGC	GAGC	CGCT	GTGC	CA (	CTGCT	rGCGCC	60 120
	ACAG	ATTT	GT C	ACC	GCG	CG G	CTTT.	rgtc <i>i</i>	A GCT	TAC:	rccg	GCC	<b>LAAA</b>	LAG A	ACTO	EAGGGG ECACCT	180
	CTGG	AGCG	GA C	TTA:	TTTA(	CC A	AGCA!	rtgg <i>i</i>	A GGI	\ATA'	rcgt	AGGT	raaa?	ATC Met	G CCI	T ATT	237
45														1			
	GGA	TCC	AAA	GAG	AGG	CCA	ACA	TTT	TTT	GAA	ATT	TTT	AAG	ACA	CGC	TGC	285
	Gly	Ser 5	Lys	Glu	Arg	Pro	Thr 10	Phe	Phe	Glu	Ile	Phe 15	ГÀв	Thr	Arg	Сув	
50	220	אאא	CCA	CAT	ጥሞእ	CCA	CCA	ΔΤΔ	ΔСТ	ىلىك	דממ	тсс	ттт	GAA	GAA	CTT	333
	Asn	Lys	Ala	Asp	Leu	Gly	Pro	Ile	Ser	Leu	Asn	Trp	Phe	Glu	Glu	Leu	
	20					25					30					35	
55	TCT	TCA	GAA	GCT	CCA	CCC	TAT	AAT	TCT	GAA	CCT	GCA	GAA Glu	GAA Glu	TCT Ser	GAA Glu	381
	ser	oer	GIU	WIG	40	110	- 7 -	nou	551	45					50		
	CAT	AAA	AAC	AAC	AAT	TAC	GAA	CCA	AAC	CTA	TTT	AAA	ACT	CCA	CAA	AGG	429
60	His	Lys	Asn	Asn 55	Asn	Tyr	Glu	Pro	Asn 60	Leu	Phe	Lys	Thr	Pro 65	Gln	Arg	

5	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	TAA Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
J	CAA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	GAT Asp	525
10	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	AAT Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT His	AAA Lys 115	573
15	AGT Ser	CTT Leu	CGC Arg	ACA Thr	GTG Val 120	AAA Lys	ACT Thr	AAA Lys	ATG Met	GAT Asp 125	CAA Gln	GCA Ala	GAT Asp	GAT Asp	GTT Val 130	TCC Ser	621
20													GTT Val				669
25													GTA Val 160				717
23													ACA Thr				765
30	ATT Ile 180	TCT Ser	GAA Glu	AGT Ser	CTA Leu	GGA Gly 185	GCT Ala	GAG Glu	GTG Val	GAT Asp	CCT Pro 190	GAT Asp	ATG Met	TCT Ser	TGG Trp	TCA Ser 195	813
35													GTG Val				861
40	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
45	AAT Asn	GTG Val	AAA Lys 230	AGC Ser	TAT Tyr	TTT Phe	TCC Ser	AAT Asn 235	CAT His	GAT Asp	GAA Glu	AGT Ser	CTG Leu 240	AAG Lys	AAA Lys	AAT Asn	957
45													ACA Thr				1005
50							Phe						AAT Asn				1053
55													CCA Pro				1101
60													GAA Glu				1149
	TTT	TCA	TTA	TGT	TTT	TCT	AAA	TGT	AGA	ACA	AAA	AAT	CTA	CAA	AAA	GTA	1197

	Phe	Ser	Leu 310	Сув	Phe	Ser	Lys	Сув 315	Arg	Thr	Lys	Asn	Leu 320	Gln	Lys	Val	
5												GAA Glu 335				GAT Asp	1245
10	_		_									AAA Lys				GTA Val 355	1293
15												GAT Asp					1341
20												ATC Ile					1389
20												ACC Thr					1437
25												CAT His 415					1485
30												ACA Thr					1533
35												CCA Pro					1581
40												ACA Thr					1629
40												GAC Asp					1077
45	Val											GCT Ala 495					1725
50												TCA Ser					1773
55												CCA Pro					1821
60												CAT His					1869
60												GAT Asp					1917

CCA GCC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA Pro Ala Thr Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT Ile Ser Thr Leu Lys Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr TGT TCT AAT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Cys His 

CCA GTA CAA CAC TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr 

5	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
10	GAT Asp 820	GTA Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
10				TAC Tyr													2781
15	AAC Asn	CAA Gln	AAC Asn	ACA Thr 855	AAT Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile 860	CAA Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT Thr	2829
20				TCA Ser													2877
25				GAG Glu													2925
30				TTA Leu													2973
				GAA Glu													3021
35				GAT Asp 935													3069
40				CTT Leu													3117
45				CTA Leu													3165
50		Lys		CCA Pro													3213
				CCA Pro					Ser					Phe			3261
55			naA	AAG Lys 1015				Leu					Ile				3309
60		Met		TTC Phe			Ile					Pro					3357

5	TGT GTT Cys Val 1045	GAA Glu	ATT Ile	GTA Val	naA	ACC Thr .050	TTG Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
5	AGC AAG Ser Lys 1060	CCT Pro	CAG Gln	Ser	ATT Ile 1065	AAT Asn	ACT Thr	GTA Val	Ser	GCA Ala L070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
10	GTA GTT Val Val	GTT Val	Ser	GAT Asp 080	TGT Cys	AAA Lys	AAT Asn	Ser	CAT His	ATA Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	TTA Leu	3501
15	TTT TCC Phe Ser	ГÀз					Ser					Thr				3549
20		Glu 1110	Ile	Thr	Glu	Leu I	Ser 1115	Thr	Ile	Leu	Glu 1	Glu 120	Ser	Gly	Ser	3597
25	CAG TTT Gln Phe 1125	Glu	Phe	Thr	Gln 1	Phe 130	Arg	Lys	Pro	Ser	Tyr 1135	Ile	Leu	Gln	Lys	3645
	AGT ACA Ser Thr 1140	Phe	Glu	Val	Pro 1145	Glu	Asn	Gln	Met 1	Thr 1150	Ile	Leu	Lys	Thr 1	Thr 155	3693
30	TCT GAG Ser Glu	Glu	Cys 1	Arg 1160	Asp	Ala	Asp	Leu I	His 1165	Val	Ile	Met	Asn 1	Ala 1170	Pro	3741
35	TCG ATT Ser Ile	Gly	Gln 175	Val	Asp	Ser	Ser	Lys 180	Gln	Phe	Glu	Gly	Thr 185	Val	Glu	3789
40		Arg 1190	Lys	Phe	Ala	Gly 1	Leu 195	Leu	Lys	Asn	Asp 1	Cys 1200	Asn	Lys	Ser	3837
45	GCT TCT Ala Ser 1205	Gly	TAT Tyr	Leu	ACA Thr	Asp	Glu	TAA naA	GAA Glu	Val	GGG Gly L215	TTT	AGG Arg	GGC Gly	TTT Phe	3885
	TAT TCT Tyr Ser 1220			Gly					Val					Leu		3933
50	AAA GCT Lys Ala		Lys					Ile					Glu			3981
55	TCT GCA Ser Ala	Glu					Ser					Lys				4029
60	TCT GTT Ser Val					Lys					Asn					4077
	AGT GAA	AAA	AAT	TAA	AAI.	TGC	CAA	CTG	ATA	TTA	CAA	AAT	AAT	ATT	GAA	4125

	Ser Glu 1285		Asn	Asn		Cys .290	Gln	Leu	Ile		Gln 1295	Asn	Asn	Ile	Glu	
5	ATG ACT Met Thr 1300	ACT Thr	GGC Gly	Thr	TTT Phe 1305	GTT Val	GAA Glu	GAA Glu	Ile	ACT Thr 1310	GAA Glu	AAT Asn	TAC	Lys	AGA Arg 1315	4173
10	AAT ACT Asn Thr	GAA Glu	Asn	GAA Glu L320	GAT Asp	AAC Asn	AAA Lys	Tyr	ACT Thr 1325	GCT Ala	GCC Ala	AGT Ser	Arg	TAA Asn 1330	TCT Ser	4221
15	CAT AAC His Asn	Leu	GAA Glu 1335	TTT Phe	GAT Asp	GGC Gly	Ser	GAT Asp 1340	TCA Ser	AGT Ser	AAA Lys	Asn	GAT Asp 1345	ACT Thr	GTT Val	4269
20	TGF ATT Cys Ile	CAT His 1350	AAA Lys	Asp GAT	GAA Glu	Thr	GAC Asp 1355	TTG Leu	CTA Leu	TTT Phe	Thr	GAT Asp 1360	CAG Gln	CAC His	AAC Asn	4317
20	ATA TGT Ile Cys 1365	Leu	AAA Lys	TTA Leu	Ser	GGC Gly L370	CAG Gln	TTT Phe	ATG Met	Lys	GAG Glu 1375	GGA Gly	AAC Asn	ACT Thr	CAG Gln	4365
25	ATT AAA Ile Lys 1380	GAA Glu	GAT Asp	Leu	TCA Ser 1385	GAT Asp	TTA Leu	ACT Thr	Phe	TTG Leu L390	GAA Glu	GTT Val	GCG Ala	Lys	GCT Ala 1395	4413
30	CAA GAA Gln Glu	GCA Ala	Cys	CAT His 1400	GGT Gly	AAT Asn	ACT Thr	Ser	AAT Asn 1405	AAA Lys	GAA Glu	CAG Gln	Leu	ACT Thr 1410	GCT Ala	4461
35	ACT AAA Thr Lys	Thr	GAG Glu 1415	CAA Gln	AAT Asn	ATA Ile	Lys	GAT Asp 1420	TTT Phe	GAG Glu	ACT Thr	Ser	GAT Asp 1425	ACA Thr	TTT Phe	4509
4.0	TTT CAG Phe Gln	ACT Thr 1430	GCA Ala	AGT Ser	GGG Gly	Lys	AAT Asn 1435	ATT Ile	AGT Ser	GTC Val	Ala	AAA Lys 1440	GAG Glu	TCA Ser	TTT Phe	4557
40	AAT AAA Asn Lys 1445	Ile	GTA Val	AAT Asn	Phe	TTT Phe 1450	GAT Asp	CAG Gln	AAA Lys	Pro	GAA Glu 1455	GAA Glu	TTG Leu	CAT His	AAC Asn	4605
45	TTT TCC Phe Ser 1460	TTA Leu	AAT Asn	Ser	GAA Glu 1465	TTA Leu	CAT His	TCT Ser	Asp	ATA Ile 1470	AGA Arg	AAG Lys	AAC Asn	Lys	ATG Met L475	4653
50	GAC ATT Asp Ile		Ser					Asp					Lys			4701
55	AAA GAA Lys Glu	A AGT 1 Ser	GTC Val	Pro	GTT Val	GGT Gly	Thr	GGA Gly 1500	AAT Asn	CAA Gln	CTA Leu	Val	ACC Thr 1505	TTC Phe	CAG Gln	4749
~~	GGA CAZ Gly Glr	A CCC A Pro 1510	Glu	. CGT . Arg	GAT Asp	Glu	AAG Lys 1515	Ile	AAA Lys	GAA Glu	Pro	ACT Thr 1520	CTG Leu	TTG Leu	GGT Gly	4797
60	TTT CAT	C ACA	GCT Ala	'AGC Ser	GGG Gly	AAA Lys	AAA Lys	GTT Val	AAA Lys	ATT	GCA Ala	AAG Lys	GAA Glu	TCT Ser	TTG Leu	4845

1525 1530 1535

5	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser 1540 1545 1550	
10	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg 1560 1565 1570	g Glu
15	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr 1575 1580 1585	
	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys 1590 1595 1600	AAC 5037 Asn
20	CTT GTT TCTTT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp 1605 1610 1615	
25	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe 1620 1625 1630	
30	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser 1640 1645 1650	Pro
35	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn 1655 1660 1665	
	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser 1670 1675 1680	Gln
40	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe 1685 1690 1695	Asp
45	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr 1700 1705 1710	
50	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn 1720 1725 1730	His
55	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser 1735 1740 1745	
	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr 1750 1760	
60	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn 1765 1770 1775	

5	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
10	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
15	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
20	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
25	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
30	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
35	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
40	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
45	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1965 1970	6141
50	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
55	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
60	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285

		-
	GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035	6333
5	CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050	6381
10	TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065	6429
15	ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT  Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe  2070 2075 2080	6477
20	GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095	6525
25	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
25	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
30	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
35	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2160	6717
40	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
45	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
50	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2225	6909
55	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
60	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG	7053

	WO 33/03/204	-
	Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	
5	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
10	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2300 2305	7149
15	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
20	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2355	7293
25	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
30	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
35	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
40	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
45	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
50	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT Asn Glu Ile His Gln Pho Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
55	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
60	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr	7773

	WO 33/03104	
	2500 2505	2510 2515
5	TCC ACT CTG CCT CGA ATC TCT CTG Ser Thr Leu Pro Arg Ile Ser Leu 2520	AAA GCA GCA GTA GGA GGC CAA GTT 7821 Lys Ala Ala Val Gly Gly Gln Val 2525 2530
10	4535	2540 2545
	CAT TGC ATA AAA ATT AAC AGC AAA His Cys Ile Lys Ile Asn Ser Lys 2550 2555	AAT GCA GAG TCT TTT CAG TTT CAC 7917 Asn Ala Glu Ser Phe Gln Phe His 2560
15	Thr Glu Asp Tyr Phe Gly Lys Glu 2565 2570	AGT TTA TGG ACT GGA AAA GGA ATA 7965 Ser Leu Trp Thr Gly Lys Gly Ile 2575
20	2580 2585	2590 2595
25	GGA AAA GAA GAA TTT TAT AGG GCT Gly Lys Glu Glu Phe Tyr Arg Ala 2600	CTG TGT GAC ACT CCA GGT GTG GAT 8061 Leu Cys Asp Thr Pro Gly Val Asp 2605 2610
30	Pro Lys Leu Ile Ser Arg Ile Trp	GTT TAT AAT CAC TAT AGA TGG ATC 8109 Val Tyr Asn His Tyr Arg Trp Ile 2620 2625
	ATA TGG AAA CTG GCA GCT ATG GAA Ile Trp Lys Leu Ala Ala Met Glu 2630 2635	A TGT GCC TTT CCT AAG GAA TTT GCT 8157 1 Cys Ala Phe Pro Lys Glu Phe Ala 2640
35	Asn Arg Cys Leu Ser Pro Glu Arg 2645 2650	G GTG CTT CTT CAA CTA AAA TAC AGA 8205 G Val Leu Leu Gln Leu Lys Tyr Arg 2655
40	TAT GAT ACG GAA ATT GAT AGA AGC Tyr Asp Thr Glu Ile Asp Arg Ser 2660	C AGA AGA TCG GCT ATA AAA AAG ATA 8253 r Arg Arg Ser Ala Ile Lys Lys Ile 2670 2675
45	ATG GAA AGG GAT GAC ACA GCT GCA Met Glu Arg Asp Asp Thr Ala Ala 2680	A AAA ACA CTT GTT CTC TGT GTT TCT 8301 a Lys Thr Leu Val Leu Cys Val Ser 2685 2690
50	GAC ATA ATT TCA TTG AGC GCA AAT Asp Ile Ile Ser Leu Ser Ala Ass 2695	T ATA TCT GAA ACT TCT AGC AAT AAA 8349 n Ile Ser Glu Thr Ser Ser Asn Lys 2700 2705
	ACT AGT AGT GCA GAT ACC CAA AA Thr Ser Ser Ala Asp Thr Gln Ly 2710 271	A GTG GCC ATT ATT GAA CTT ACA GAT 8397 'B Val Ala Ile Ile Glu Leu Thr Asp .5 2720
55	GGG TGG TAT GCT GTT AAG GCC CA Gly Trp Tyr Ala Val Lys Ala Gl 2725 2730	AG TTA GAT CCT CCC CTC TTA GCT GTC 8445 In Leu Asp Pro Pro Leu Leu Ala Val 2735
60	TTA AAG AAT GGC AGA CTG ACA GT Leu Lys Asn Gly Arg Leu Thr Va 2745	TT GGT CAG AAG ATT ATT CTT CAT GGA 8493 al Gly Gln Lys Ile Ile Leu His Gly 2750 2755

5	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
10	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
15	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
20	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835	8733
25	TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala 2840 2845 2850	8781
	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 ,2865	8829
30	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
35	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
40	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala 2900 2905 2910 2915	8973
45	TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn 2920 2925 2930	9021
	CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 2940 2945	9069
50	ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 2955 2960	9117
55	AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 2970 2975	9165
60	AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Sei Asp 2980 2985 2990 2995	9213

	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
5	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
10	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
15	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
20	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA  Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile  3060 3065 3070 3075	9453
	GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501
25	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
30	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120	9597
35	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
40	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150	9693
	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
45	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
50	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
55	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
60	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG	9981

	Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	•
5	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys 3255 3260 3265	10029
10	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
15	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
•	CCA CCA AGG AGT TCT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 3305 3310 3315	10173
20	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	10221
25	TCT CTT TTG GAA AGT AAT TCA ATA GCI GAC GAA GAA CTT GCA TTG ATA Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	10269
30	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3355 3360	10317
35	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	10365
	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	10413
40	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	10461
45	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
50	(2) INFORMATION FOR SEQ ID NO:9:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 3418 amino acids</li> <li>(B) TYPE: amino acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	<ul><li>(ii) MOLECULE TYPE: protein</li><li>(v) FRAGMENT TYPE: internal</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:</li></ul>	

			Ile													
-			Сув	~ ~	Lys				<i>_</i>					<b>-</b>		
5			Leu 35					A()					<b>T</b> J			
			Glu				55					00				
10			Arg			Ser	Tyr				10					
	Phe		Glu		0 5					90						
15			Asp	1 00					105					T T O		
13			Lys 115					1.20					143			
			Ser				125					140				
20	- 4-		Gln			150					122					100
	Cys		Ser		165					1/0					<b>_</b> , _	
25			His	100					185					100		
20			Ser 195					200					203			
		210	Val				215					220				
30	225		Ala			230					235					240
	_		Asn		245					250					233	
35			Arg	200					265					210		
			Lys 275					- 2 KD					200			
		200	Leu				295					300				
40	205	-	Ser			3 1 ก					3 1 2					520
			Val		つつに					330	)				333	
45			Asp	340					345	•				330		
			val 355					360	)				200	1		
		271	•				375	5				300				Ser
50	20	_				3 90					395	)				Leu 400 Ile
					405					410	j				41-	Ile Glu
55				421	١				42	5				43	,	
			421	_				44	U				77.	,		Arg Val
		4 5	^				45	5				40	,			Cys 480
60		-				471	`				42 / 3	_				480 a Ser
	Il	e Le	u Al	a va	т гу	2 611	, MI	.u .i.	- 56	_ 01	,					

	Com	Dha	aln	alv	485 Tle	Taya	Lvg	Ser	Tle	490 Phe	Ara	Ile	Arq	Glu	495 Ser	Pro
_				500					505					510		
5	-		515					520					525	Asp		
		530					535					540		Ile		
10	545	_				550					555			Ile		560
	_				565					570				Leu	575	
	Ala	Gly	Leu	Ile 580	Ser	Thr	Leu	Lys	Lys 585	Lys	Thr	Asn	Lys	Phe 590	Ile	Tyr
15	Ala	Ile	His 595		Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	ГÀЗ	Asp
	Gln	Lys 610	Ser	Glu	Leu	Ile	Asn 615	Сув	Ser	Ala	Gln	Phe 620	Glu	Ala	Asn	Ala
20	Phe 625	Glu	Ala	Pro	Leu	Thr 630		Ala	Asn	Ala	^sp 635	Ser	Gly	Leu	Leu	His 640
20	Ser	Ser	Val	Lys	Arg 645	Ser	Cys	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr
				660	Ser				665					Cys 670		
25	Asn	Glu	Thr 675	Cys	Ser	Asn	Asn	Thr 680	Val	Ile	Ser	Gln	Asp 685	Leu	Asp	Tyr
	Lys	Glu 690	Ala	Lys	Cys	Asn	Lys 695	Glu	Lys	Leu	Gln	Leu 700	Phe	Ile	Thr	Pro
30	705	Ala	_			710					715			Glu		720
30	Pro				725					730				Leu	735	
	Ala	Сув	His	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	Asp 750	Thr	Asp
35			755					760					765	Ala		
		770					775					780		Leu		
40	785					790					795			Leu		800
					805					810				Pro	815	
	_			820					825					Asn 830		
45			835					840					845			Lys
		850					855					860		Lys		
50	865					870					875					Glu 880
					885					890				Val	895	
				900					905					His 910		
55			915					920					925			
		930	)				935					940		Ser		
60	945					950					955					<b>Lys</b> 960
	Gln	His	: Ile	. Lys	Met 965		Lev	Gly	Gln	970		Lys	Ser	Asp	11e 975	Ser

				980					985					Met 990		
5	_		995					100	0				100			
		101	0				101	5		_		1020	)	His		
	1029	5		_		1030	)	_	_		103	5		Tyr		104
10				_	104	5				1050	כ			Asp	105	5
	_	_		1060	ס ֿ				106	5				Ala 1070	)	
15			1075	5				1080	o ¯	_			108			
		1090	כ				109	5				1100	)	Asn		
	Pro 1109		GIn	rys	Ala	G1u 1110		Thr	GIu	Leu	Ser		Ile	Leu	GIu	G1u 112
20			Ser	Gln	Phe 1125	Glu		Thr	Gln	Phe 1130	Arg		Pro	Ser	Tyr 1135	Ile
	Leu	Gln	Lys	Ser 1140		Phe	Glu	Val	Pro 1145		Asn	Gln	Met	Thr 1150		Leu
25			1155	5				1160	)				1165			
		1170	כ			_	1175	5	_			1180	)	Phe		_
	1185	5				1190	)				1195	5	_	Asn	_	120
30	Asn	Lys	Ser	Ala	Ser 1205	_	Tyr	Leu		Asp 1210		Asn	Glu	Val	Gly 1215	
	_	_		1220	)			_	1225	5				Ser 1230	)	
35	Ala	Leu	Gln 1235		Ala	Val	Lys	Leu 1240		Ser	Asp	Ile	Glu 1245	Asn	Ile	Ser
		1250	)				1255	5				1260	1	Ser		_
	1265	5				1270	)			-	1275			His		128
40	_				1285	5			-	1290	)				1295	
	Asn	Ile	Glu	Met 1300		Thr	Gly	Thr	Phe 1305		Glu	Glu	Ile	Thr 1310		Asn
45	Tyr		Arg 1315	Asn		Glu		Glu 1320	Asp		Lys		Thr 1325	Ala		Ser
	Arg		Ser		Asn	Leu		Phe		Gly	Ser		Ser	Ser	Lys	Asn
	Asp 1345		Val	Сув	Ile	His		Asp	Glu	Thr	Asp 1355		Leu	Phe	Thr	Asp 136
50	Gln	His	Asn	Ile	Cys 1365		Lys	Leu	Ser	Gly 1370		Phe	Met	Lys	Glu 1375	
	Asn	Thr	Gln	Ile 1380	_	Glu	Asp	Leu	Ser 1385	_	Leu	Thr	Phe	Leu 1390		Val
55	Ala	Lys	Ala 1395		Glu	Ala	Сув	His 1400	_	Asn	Thr	Ser	Asn 1405	Lys	Glu	Gln
	Leu	Thr 1410		Thr	Lys	Thr	Glu 1415		Asn	Ile	Lys	Asp 1420		Glu	Thr	Ser
	Asp 1425		Phe	Phe	Gln	Thr 1430		Ser	Gly	ГÀЗ	Asn 1435		Ser	Val	Ala	Lys 144
60	Glu	Ser	Phe	Asn	Lys 1445		Val	Asn	Phe	Phe 1450	Asp		Lys	Pro	Glu 1455	
	Leu	His	Asn	Phe	Ser	Leu	Asn	Ser	Glu			Ser	Asp	Ile	Arg	Lys

				146					146					147		
	Asn	Lys	Met 147		Ile	Leu	Ser	Tyr 148		Glu	Thr	Asp	Ile 148		Lys	Hi
5	Lys	Ile 149	Leu 0	Lys	Glu	Ser	Val 149		Val	Gly	Thr	Gly 150		Gln	. Leu	va:
	Thr 150		Gln	Gly	Gln	Pro 151		Arg	Asp	Glu	Lys 151		Lys	Glu	Pro	Th:
10			Gly		152	5				153	0				153	5
			Leu	154	0				154	5				155	0	
			Glu 155	5				156	0				156	5		_
15		157	-				157	5				158	0			
	158	5	Ala			159	0				159	5				160
20			Asn		160	5				161	0			_	161	5
			Asn	162	0				162	5				163	0	
			Leu 163!	5				164	0				164	5		
25		165					165	5				1660	0			
	1665	5	Ser			1670	)				1679	5				168
30			Gln		1685	5				1690	)				169	5
			Asp	1700	)				1709	5				1710	)	
			Leu 1715	5				1720	)				1725	5		
35		1730					1735	5				1740	)			
	1745	5	Asn			1750	)				1755	5	_			176
40			Leu		1765	;				1770	)				1775	5
			Val	1780	)				1785	5				1790	)	
			Lys 1795	5				1800	)				1805	5		
45		1810					1815	5				1820	)			
	1825	5	Ile			1830	)				1835	;				184
50			Ala		1845	5				1850	)				1855	5
			Ile	1860	)				1865	5				1870	)	
			Lys 1875	5				1880	)				1885	<b>,</b>		
55		1890					1899	;				1900	)			
	H18		Ser	Leu	Asp	Asn 1910		Glu	Сув	Ser	Thr 1915		Ser	His	Lys	Val 192
60			Asp		1925	;				1930	Gln	His			1935	Met
	Ser	Gly	Leu	Glu 1940										Val		Leu

			195	5		_		1960	0				196	5		Ser
5	Val	Ser 197		Ala	Asn	Thr	Cys 197		Ile	Phe	Ser	Thr 198		Ser	Gly	Lys
	Ser 198		Gln	Val	Ser	Asp 199		Ser	Leu	Gln	Asn 199		Arg	Gln	Val	Phe 200
					2009	5		-		201	0		_		201	
10	-			Glu 2020	)		_		202	5	-			203	0	
		_	203					2040	0				204	5		
15		205	0	Ser			205	5				206	0			
	206	5				2070	כ				207	5				Leu 208
				Asp	2085	5				2090	0				209	5
20			_	2100	)			_	210	5		_		2110	)	Arg
			211		-			2120	)			_	212	5		_
25		2130	o -				213	5				214	۔ ت			Glu
	2149	5				2150	)			-	215	5				Gln 216
		_		Gln	2165	5				2170	)				2175	5
30				Leu 2180	)	_			2185	5		_		2190	)	
			219					2200	)				2209	5		
35		2210	)	Сув			2219	5	-			2220	)			
	2225	5		Val		2230	)				2235	5				224
		_		Lys	2245	5				2250	)				2255	;
40				Glu 2260	)				2265	5				2270	)	
	_	_	2275					2280	)				2285	5		
45		229	)	Glu			2295	5				2300	)			
	2309	5		Lys		2310	)	_			2315	5				232
				His	2325	5				2330	)				2335	5
50	Thr	Thr	Lys	Glu 2340	_	Gln	Glu	Ile	Gln 2345		Pro	Asn	Phe	Thr 2350		Pro
	Gly	Gln	Glu 235	Phe 5	Leu	Ser	Lys	Ser 2360		Leu	Tyr	Glu	His 2369		Thr	Leu
55	Glu	Lys 237		Ser	Ser	Asn	Leu 237		Val	Ser	Gly	His 238		Phe	Tyr	Gln
	238	5		Thr	_	2390	)				2399	5				240
	Arg	Pro	Thr	Lys	Val 240		Val	Pro	Pro	Phe 2410		Thr	Lys	Ser	His 2415	
60		_		Glu 2420	)	-		_	242	5				2430	)	_
	Gln	ràs	Gln	Asn	Ile	Asp	Gly	His	Gly	Ser	Asp	Asp	Ser	Lys	Asn	Lys

			243:	•				244	,				444	<b>-</b>		
	Ile	Asn 2450	_	Asn	Glu	Ile	His 245		Phe	Asn	Lys	Asn 246		Ser	Asn	Glr
5	Ala 246	Ala		Val	Thr	Phe 2470		Lys	Сув	Glu	Glu 247		Pro	Leu	Asp	Let 248
			Co-	Tan	Cln			Ara	Acn	Tle			Mat	Arg	Tla	
					2485	5				249	0				249	5
10	Lys	ГÀЗ	Gln	Arg 250		Arg	Val	Phe	Pro 250		Pro	Gly	Ser	Leu 251		Leu
	Ala	Lys	Thr 2519		Thr	Leu		Arg 2520		Ser	Leu	Lys	Ala 252	Ala 5	Val	Gly
	Gly	Gln 2530		Pro	Ser	Ala	Cys 2535		His	Lys	Gln	Leu 2540		Thr	Tyr	Gly
15	Val 2549		Lys	His	Cys	Ile 2550	_	Ile	Asn	Ser	Lys 2559		Ala	Glu	Ser	Phe 256
			His	Thr	Glu 2569	-	Tyr	Phe	Gly	Lys 2570		Ser	Leu	Trp	Thr 2579	
20	Lys	Gly	Ile	Gln 2580		Ala	Asp	Gly	Gly 2589		Leu	Ile	'Pro	Ser 2590		Asp
	Gly	Lys	Ala 2599	_	Lys	Glu	Glu	Phe 2600		Arg	Ala	Leu	Cys 2605	Asp	Thr	Pro
	Gly	Val 2610		Pro	Lys	Leu	Ile 2615		Arg	Ile	Trp	Val 2620		Asn	His	Tyr
25	Arg 2629	_	Ile	Ile	_	Lys 2630		Ala	Ala	Met	Glu 2635	_	Ala	Phe	Pro	Lys 264
	Glu	Phe	Ala	Asn	Ara	Cvs	Leu	Ser	Pro	Glu	Ara	Val	Leu	Leu	Gln	Leu
					2645	5				2650	)			Ser	2655	5
30				2660	)				2665	5				2670	)	
			2675	5				2680	)				2685			
2.5		2690	)				2695	;				2700	)	Glu		
35	2705	5				2710	)				2715	5		Ile		272
			_	_	2725	5			_	2730	)			Pro	2735	5
40				2740	)			_	2745	5				Lys 2750	)	
			2755	5				2760	)		_		2765			
4.5		2770	)				2775	5				2780	)	Ser		
45	Pro 2785	5				2790	)				2795	5				280
					2805	5				2810	)			Asn	2815	5
50				2820	)				2825	5				Trp 2830	)	
			2835	5				2840	)				2845			
	-	2850	)		_		2855	5				2860	)	Leu		
55	2865	5		-		2870	)				2875	5		Glu		288
		_		•	2889	5				2890	)			Gln	2895	5
60				2900	)				2905	5				Asn 2910	)	
	Asp	Pro	Ala 291		Leu	Glu		Tyr		Ser	Glu		Gln 2929	Leu	Arg	Ala

	БСи	2930		HIP	Arg	GIII	293		7011	mp	בעם	294		AIU	0211	
5	Gln 2945		Glu	Ile	Arg	Lys 295		Met	Glu	Ser	Ala 295		Gln	Lys	Glu	Gln 296
	Gly	Leu	Ser	Arg	Asp 2969	Val		Thr	Val	Trp 297		Leu	Arg	Ile	Val 297!	
	Tyr	Ser	Lys	Lys 2980	Glu	Lys	Asp	Ser	Val 298		Leu	Ser	Ile	Trp 299	_	Pro
10	Ser	Ser	Asp 2999	Leu	Tyr	Ser	Leu	Leu 3000	Thr		Gly	Lys	Arg 300	Tyr		Ile
	Tyr	His 3010	Leu	-	Thr	Ser	Lys 301	Ser		Ser	Lys	Ser 302	Glu		Ala	Asn
15	Ile 3025	Gln		Ala	Ala	Thr 3030	Lys		Thr	Gln	Tyr 303	Gln		Leu	Pro	Val
			Glu	Ile	Leu 3049	Phe		Ile	Tyr	Gln 3050	Pro		Glu	Pro	Leu 3059	
	Phe	ser	Lys	Phe 3060	Leu		Pro	Asp	Phe 3065		Pro	Ser	Cys	Ser 3070	Glu	
20	qaA	Leu	Ile 3079	Gly	Phe	Val	Val	Ser 3080	Val		Lys	Lys	Thr 3089	Gly		Ala
	Pro	Phe 3090	Val	_	Leu	Ser	Asp 3099	Glu		Tyr		Leu 3100	Leu		Ile	Lys
25	Phe 3105	-	Ile	Asp	Leu	Asn 3110	Glu		Ile	Ile		Pro		Met	Leu	Ile 312
	Ala	Ala	Ser	Asn	Leu 3125		Trp	Arg	Pro	Glu 3130		Lys	Ser	Gly	Leu 3135	
	Thr	Leu	Phe	Ala 3140	Gly	Asp	Phe	Ser	Val 3145		Ser	Ala	Ser	Pro 3150	_	Glu
30	Gly	His	Phe 3155		Glu	Thr	Phe	Asn 3160		Met	ГÀЗ	Asn	Thr 3165		Glu	Asn
	Ile	Asp 3170		Leu	Cys	Asn	Glu 3175		Glu	naA	Lys	Leu 3180		His	Ile	Leu
35	His 3185		Asn	Asp	Pro	Lys 3190		Ser	Thr	Pro	Thr 3195		Asp	Сув	Thr	Ser 320
	_		-		Ala 3205	5				3210	)	_		_	3215	5
				3220		_			3225	5				3230	)	
			3235	;	Arg	_		3240	)				3245	i		
		3250	)		Сув		3255	<b>,</b>				3260	)			
	3265	;			Arg	3270	)				3275	i				328
					Pro 3285	,				3290	)				3295	5
				3300				_	3305	5	_			3310	)	
	-	_	3315	5	Leu			3320	)				3325	; -	-	
		3330	)		Leu		3335	;				3340	)			
55	3345	;			Thr	3350	)				3355	i				336
					Val 3365	i				3370	)				3375	5
		_	=	3380			-	_	3385	5				3390	)	_
60			3395	5	Ser			3400	)	•	GIU	cys	Glu 3405		ASN	гув
	Gln	Asp	Thr	Ile	Thr	Thr	Lys	Lys	Tyr	Ile						

3410 3415

5			(2)	INI	FORM	ATIO	v FOI	R SE	Q ID	NO:	10:						
		(:	(A) (B)	EQUEN LENC TYPI STRA	FTH: E: nu	1048 acle:	85 ba	ase p	pair	5							
10		•	. ,	TOPO													
		•	•	MOLE FEAT		TYPI	E: Ci	JNA									
15			(B)	) NAM LOC ) OTI	CATIO	ON: 3	229.	104	182		OMI4	)					
20		•		EQUI													
	TCT	GCTG(	CGC (	CTCG(	GTGT GCGC	rc Ti	TTTG( TTTT	CGGC(	GTC	GGT(	CCC4	GCC	GGA(	GAA (	GCGT(	TGCGCC GAGGGG GCACCT F ATT	60 120 180 237
25	C10.															o Ile	
30				GAG Glu													285
35				GAT Asp													333
33				GCT Ala													381
40				AAC Asn 55													429
45				TAT Tyr													477
50				ACT Thr													525
55				TTA Leu													573
<i></i>				ACA Thr													621
60				CTA Leu 135													669

5	TGT Cys	ACA Thr	CAT His 150	GTA Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	GAT Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	TGT Cys	GGG Gly	AGT Ser	717
10					CCA Pro												765
10					CTA Leu												813
15	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
20					GCA Ala												909
25					TAT Tyr												957
2.0					GCT Ala												1005
30					CAT His												1053
35					AAA Lys 280												1101
40					TAT Tyr												1149
45					TTT Phe												1197
50					ACT Thr												1245
30					TCT Ser												1293
55					CCA Pro 360												1341
60					TTT Phe												1389

	GTA Val	CCG Pro	TCT Ser 390	TTG Leu	GCC Ala	TGT Cys	GAA Glu	TGG Trp 395	TCT Ser	CAA Gln	CTA Leu	ACC Thr	CTT Leu 400	TCA Ser	GGT Gly	CTA Leu	1437
5	AAT Asn	GGA Gly 405	GCC Ala	CAG Gln	ATG Met	GAG Glu	AAA Lys 410	ATA Ile	CCC Pro	CTA Leu	TTG Leu	CAT His 415	ATT Ile	TCT Ser	TCA Ser	TGT Cys	1485
10	GAC Asp 420	CAA Gln	AAT Asn	ATT Ile	TCA Ser	GAA Glu 425	AAA Lys	GAC Asp	CTA Leu	TTA Leu	GAC Asp 430	ACA Thr	GAG Glu	AAC Asn	AAA Lys	AGA Arg 435	1533
15	AAG Lys	AAA Lys	GAT Asp	TTT Phe	CTT Leu 440	ACT Thr	TCA Ser	GAG Glu	AAT Asn	TCT Ser 445	TTG Leu	CCA Pro	CGT Arg	ATT Ile	TCT Ser 450	AGC Ser	1581
20	CTA Leu	CCA Pro	AAA Lys	TCA Ser 455	GAG Glu	AAG Lys	CCA Pro	TTA Leu	AAT Asn 460	GAG Glu	GAA Glu	ACA Thr	GTG Val	GTA Val 465	AAT Asn	AAG Lys	1629
0.5	AGA Arg	GAT Asp	GAA Glu 470	GAG Glu	CAG Gln	CAT His	CTT Leu	GAA Glu 475	TCT Ser	CAT His	ACA Thr	GAC Asp	TGC Cys 480	ATT Ile	CTT Leu	GCA Ala	1677
25	GTA Val	AAG Lys 485	CAG Gln	GCA Ala	ATA Ile	TCT Ser	GGA Gly 490	ACT Thr	TCT Ser	CCA Pro	GTG Val	GCT Ala 495	TCT Ser	TCA Ser	TTT Phe	CAG Gln	1725
30	GGT Gly 500	ATC Ile	AAA Lys	AAG Lys	TCT Ser	ATA Ile 505	TTC Phe	AGA Arg	ATA Ile	AGA Arg	GAA Glu 510	TCA Ser	CCT Pro	AAA Lys	GAG Glu	ACT Thr 515	1773
35	TTC Phe	AAT Asn	GCA Ala	AGT Ser	TTT Phe 520	TCA Ser	GGT Gly	CAT His	ATG Met	ACT Thr 525	GAT Asp	CCA Pro	AAC Asn	TTT Phe	AAA Lys 530	AAA Lys	1821
40	GAA Glu	ACT Thr	GAA Glu	GCC Ala 535	TCT Ser	GAA Glu	AGT Ser	GGA Gly	CTG Leu 540	GAA Glu	ATA Ile	CAT His	ACT Thr	GTT Val 545	TGC Cys	TCA Ser	1869
45	Gln	Lys	GAG Glu 550	Asp	Ser	Leu	Cys	Pro	Asn	Leu	Ile	Asp	AAT Asn 560	GGA Gly	AGC Ser	TGG Trp	1917
45	CCA Pro	GCC Ala 565	Thr	ACC Thr	ACA Thr	CAG Gln	AAT Asn 570	TCT Ser	GTA Val	GCT Ala	TTG Leu	AAG Lys 575	AAT Asn	GCA Ala	GGT Gly	TTA Leu	1965
50	ATA Ile 580	Ser	ACT Thr	TTG Leu	AAA Lys	AAG Lys 585	Lys	ACA Thr	AAT Asn	AAG Lys	TTT Phe 590	ATT	TAT Tyr	GCT Ala	ATA Ile	CAT His 595	2013
55	gat Asp	GAA	ACA Thr	TCT Ser	TAT Tyr 600	Lys	GGA Gly	AAA Lys	AAA Lys	ATA Ile 605	Pro	AAA Lys	GAC Asp	CAA Gln	AAA Lys 610	TCA Ser	2061
60	GAA Glu	CTA Leu	ATT	AAC Asn 615	Cys	TCA Ser	GCC Ala	CAG Gln	Phe 620	Glu	GCA Ala	AAT Asn	GCT Ala	TTT Phe 625	GAA Glu	GCA Ala	2109
	CCF	CT	r aca	TTT	GCA	raa .	GCI	GAT	TCA	GGI	TTA	TTG	CAT	TCT	TCT	GTG	2157

	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640	Ser	Ser	Val	
5	AAA Lys	AGA Arg 645	AGC Ser	TGT Cys	TCA Ser	CAG Gln	AAT Asn 650	GAT Asp	TCT Ser	GAA Glu	GAA Glu	CCA Pro 655	ACT Thr	TTG Leu	TCC Ser	TTA Leu	2205
10	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
15	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
0.0	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
20	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
25	FÅ2 YYY	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAĞ Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
30	CCA Pro 740	GTA Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
35	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
	ACT Thr	CCT Pro	ACT Thr	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	AAC Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
40	GGC Gly	AAA Lys	GAA Glu 790	TCA Ser	TAC Tyr	AA. Lys	ATG Met	TCA Ser 795	GAC Asp	AAG Lys	CTC Leu	AAA Lys	GGT Gly 800	AAC Asn	AAT Asn	TAT Tyr	2637
45	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	Glu	AAG Lys	AAT Asn	CAA Gln	2685
50	GAT Asp 820	Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	Glu	AAT Asn	TAT Tyr	Lys	AAC Asn 830	Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
55	CCT Pro	GAA Glu	AAA Lys	TAC Tyr	ATG Met 840	Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC Phe	2781
	AAC Asn	CA/	A AAC	ACA Thr	Aen	CTA	AGA Arg	GTA Val	ATC Ile	Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	Glu	ACT Thr	2829
60	ACT Thr	TC	A ATT	TCF Ser	AAA Lys	ATA	ACT Thr	GTC Val	CAA C	CCA Pro	GAC Asp	TCT Ser	GAA	GAA Glu	CTT Leu	TTC Phe	2877

PCT/US98/16905 WO 99/09164

	WO 99/091	64												P	L 1/US9	8/10302
		870					875					880				
5	TCA GAC Ser Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
10	AAT CTT Asn Leu 900	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
3.5	TGT GTA Cys Val	AAC Asn	GAA Glu	CCC Pro 920	ATT Ile	TTC Phe	AAG Lys	AAC Asn	TCT Ser 925	ACC Thr	ATG Met	GTT Val	TTA Leu	TAT Tyr 930	GGA Gly	3021
15	GAC ACA Asp Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT Ile	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
20	GTT TAT Val Tyr	GTT Val 950	CTT Leu	GCA Ala	GAG Glu	GAG Glu	AAC Asn 955	AAA Lys	AAT Asn	AGT Ser	GTA Val	AAG Lys 960	CAG Gln	CAT His	ATA Ile	3117
25	AAA ATG Lys Met 965	Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	AAT Asn	ATA Ile	3165
30	GAT AAA Asp Lys 980	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	AAC Asn	rys Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
0.5	CTC TTA Leu Leu	GGT	Pro	ATT Ile 1000	Ser	TAA Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg 1010	ACA Thr	3261
35	GCT TCA Ala Ser	Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser 1020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
40	AAA ATG Lys Met	TTC Phe 1030	TTC Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr 1040	AGT Ser	TTA Leu	GCT Ala	3357
45	TGT GTT Cys Val 1045	Glu	ATT Ile	GTA Val	Asn	ACC Thr 1050	Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
50	AGC AAC Ser Lys 1060	CCT Pro	CAG Gln	Ser	ATT Ile 1065	.» sn	ACT Thr	GTA Val	Ser	GCA Ala 1070	His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
	GTA GTT Val Val	r GTT l Val	Ser	GAT Asp 1080	Сув	Lys	AAT Asn	AGT Ser	CAT His 1085	Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	Leu	3501
55	TTT TCC	r Lys	CAG Gln 1095	Asp	TTT Phe	AAT Asn	TCA Ser	AAC Asn 1100	His	' AAT Asn	TTA Leu	Thr	CCT Pro 1105	Ser	CAA Gln	3549
60	AAG GC	A GAA a Glu 1110	Ile	ACA Thr	GAA	CTI Lev	TCT Ser 1115	Thr	T ATA	TTA Leu	GAA Glu	GAA Glu 1120	Ser	GGA Gly	AGT Ser	3597

5	CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125	3645
	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150	3693
10	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
15	TCG ATT GGT CAG GTA GAC AGC AGC CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
20	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA ATT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
25	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
30	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
35	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
40	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
45	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
50	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
55	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
60	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317

	WO 99/09104	-
_	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365
5	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT  Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala  1380 1385 1390 1395	4413
10	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
15	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe  1415 1420 1425	4509
20	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
25	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605
25	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470	4653
30	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
35	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
40	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
45	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
43	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1555	4893
50	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565	4941
55	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
60	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT	5085

		-
	Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	
5	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1635	5133
10	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
15	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
20	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
25	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1715	5373
30	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
35	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
40	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
45	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
50	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	566 <sub>1</sub>
55	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
32	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
60	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile	5805

PCT/US98/16905 WO 99/09164

	1845 1850	
5	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
10	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
15	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
20	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
25	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
30	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970	6141
	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
35	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
40	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285
45	GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035	6333
50	CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050	6381
	TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065	6429
55	ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe 2070 2075 2080	6477
60	GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095	6525

5	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
10	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
15	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
20	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
25	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
30	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
35	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
40	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
45	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile L; Arg Asn Leu Leu Asn 2280 2285 2290	7101
50	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2305 2300 2305	7149
55	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
60	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245

	MO 33/03104	
	GAA CGT CAA GAG ATA CAG AAT CCA AAT TIT ACC GCA CGT GGT GGT GGT GGT GGT GGT GGT GGT GGT	7293
5	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
10	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
15	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
20	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AG CAA AAG CAA Glu Glu Glu Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
25	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
30	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
35	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
40	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2515	7773
45	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
50	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 . 2540 2545	7869
55	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2560	7917
60	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly 1le 2565 2570 2575	7965
	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT	8013

	Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	
5	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	8061
10	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
15	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640	8157
	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2005	8205
20	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2670 2675	8253
25	ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690	8301
30	GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705	8349
35	ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720	8397
	GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 2730 2735	8445
40	TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755	8493
45	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
50	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
55	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
60	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser	8733

	2820	2825	2830	2835
5	TCT GGA TI Ser Gly Le	TA TAC ATA TTT CGC tu Tyr Ile Phe Arg 2840	AAT GAA AGA GAG GAA G Asn Glu Arg Glu Glu G 2845	AA AAG GAA GCA 8781 lu Lys Glu Ala 2850
10	GCA AAA TA Ala Lys Ty	AT GTG GAG GCC CAA /r Val Glu Ala Gln 2855	CAA AAG AGA CTA GAA G Gln Lys Arg Leu Glu A 2860	CC TTA TTC ACT 8829 Na Leu Phe Thr 2865
	AAA ATT CA Lys Ile Gl 28	ln Glu Glu Phe Glu	GAA CAT GAA GAA AAC A Glu His Glu Glu Asn 7 2875 28	ACA ACA AAA CCA 8877 Thr Thr Lys Pro 880
15	TAT TTA CO Tyr Leu Pr 2885	CA TCA CGT GCA CTA ro Ser Arg Ala Leu 2890	ACA AGA CAG CAA GTT ( Thr Arg Gln Gln Val 2 2895	CGT GCT TTG CAA 8925 Arg Ala Leu Gln
20	GAT GGT G Asp Gly A 2900	CA GAG CTT TAT GAA la Glu Leu Tyr Glu 2905	GCA GTG AAG AAT GCA ( Ala Val Lys Asn Ala ( 2910	GCA GAC CCA GCT 8973 Ala Asp Pro Ala 2915
25	TAC CTT G Tyr Leu G	AG GGT TAT TTC AGT lu Gly Tyr Phe Ser 2920	GAA GAG CAG TTA AGA G Glu Glu Gln Leu Arg 2 2925	GCC TTG AAT AAT 9021 Ala Leu Asn Asn 2930
30	CAC AGG C His Arg G	AA ATG TTG AAT GAT In Met Leu Asn Asp 2935	AAG AAA CAA GCT CAG Lys Lys Gln Ala Gln 2940	ATC CAG TTG GAA 9069 Ile Gln Leu Glu 2945
	Ile Arg L	AG GCC ATG GAA TC ys Ala Met Glu Sei 50	GCT GAA CAA AAG GAA Ala Glu Gln Lys Glu 2955 2	CAA GGT TTA TCA 9117 Gln Gly Leu Ser 960
35	AGG GAT G Arg Asp V 2965	TC ACA ACC GTG TGC al Thr Thr Val Tr 2970	AAG TTG CGT ATT GTA Lys Leu Arg Ile Val 2975	AGC TAT TCA AAA 9165 Ser Tyr Ser Lys
40	AAA GAA A Lys Glu I 2980	AAA GAT TCA GTT AT Lys Asp Ser Val Ilo 2985	A CTG AGT ATT TGG CGT be Leu Ser Ile Trp Arg 2990	CCA TCA TCA GAT 9213 Pro Ser Ser Asp 2995
45	TTA TAT T	CTT CTG TTA ACA GA Ser Leu Leu Thr Gl 3000	A GGA AAG AGA TAC AGA u Gly Lys Arg Tyr Arg 3005	ATT TAT CAT CTT 9261 Ile Tyr His Leu 3010
50	GCA ACT	TCA AAA TCT AAA AG Ser Lys Ser Lys Se 3015	T AAA TCT GAA AGA GCT r Lys Ser Glu Arg Ala 3020	AAC ATA CAG TTA 9309 Asn Ile Gln Leu 3025
	Ala Ala	ACA AAA AAA ACT .CA Thr Lys Lys Thr Gl 030	G TAT CAA CAA CTA CCG n Tyr Gln Gln Leu Pro 3035	GTT TCA GAT GAA 9357 Val Ser Asp Glu 3040
55	ATT TTA Ile Leu 3045	TTT CAG ATT TAC CA Phe Gln Ile Tyr Gl 305	G CCA CGG GAG CCC CTT n Pro Arg Glu Pro Leu 0 3055	CAC TTC AGC AAA 9405 His Phe Ser Lys
60	TTT TTA Phe Leu 3060	GAT CCA GAC TTT CA Asp Pro Asp Phe G1 3065	AG CCA TCT TGT TCT GAG In Pro Ser Cys Ser Glu 3070	GTG GAC CTA ATA 9453 Val Asp Leu Ile 3075

5	GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CGT TO GGG GGA TTT GTC GTT GTC GTT GTG AAA AAA ACA GGA CTT GCC CGT TO GCC GTT GCC	9501
	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
10	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3120	9597
15	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
20	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140	9693
25	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
30	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
35	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Met Ser Ser 3205 3210 3215	9885
40	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Ty Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
45	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
5.0	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3265	10029
50	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
55	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
60	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 3305 3310 3315	10173

	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	221
5	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA  Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile  3335 3340 3345	0269
10	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 1 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3355 3360	0317
15	Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr  3365  3370  3375	0365
20	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3385 3390 3395	0413
	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 1 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	.0461
25	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
30	(2) INFORMATION FOR SEQ ID NO:11:	
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 3418 amino acids</li> <li>(B) TYPE: amino acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	<ul><li>(ii) MOLECULE TYPE: protein</li><li>(v) FRAGMENT TYPE: internal</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:</li></ul>	
	Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys	
45	1 5 10 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe	
	Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu	
50	Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr	
	Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 70 75 80	
55	Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 90 95	
	Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser  100 105 110	
	Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp  125  125  127  128  129  129  120  125	
60	Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val	
	Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val	

	145					150					155					160
	Сув				165	His				170			Gly		175	
5		=		180					185				Asp	190		
			195					200					Ser 205			
10		210		_			215					220	Phe			
	225					230					235		Asp			240
1.5					245					250			Ser		255	
15			_	260					265				Thr Lys	270		
			275					280					285 Asp			
20		290			-		295					300	Thr			
	305	-				310	-			-	315	_	Phe	_		320
25				_	325					330			Lys		335	
				340					345				Pro	350		
			355					360					365 Asp			
30		370				_	375					380	Gln			
	385 Ser	Gly	Leu	Asn	Gly	390 Ala	Gln	Met	Glu	Lys	395 Ile	Pro	Leu	Leu	His	400 Ile
35	Ser	Ser	Cys	_	405 Gln		Ile	Ser		410 Lys	Asp	Leu	Leu		415 Thr	Glu
	Asn	Lys		420 Lys	Lys	Asp	Phe		425 Thr	Ser	Glu	Asn	Ser	430 Leu	Pro	Arg
	Ile		435 Ser	Leu	Pro	Lys		440 Glu	Lys	Pro	Leu		445 Glu	Glu	Thr	Val
40		450 Asn	Lys	Arg	Asp		455 Glu	Gln	His	Leu	Glu 475	460 Ser	His	Thr	Asp	Cys 480
	465 Ile	Leu	Ala	Val	Lys 485	470 Gln	Ala	Ile	Ser	Gly 490		Ser	Pro	Val	Ala 495	
45	Ser	Phe	Gln	Gly 500		Lys	Lys	Ser	Ile 505		Arg	Ile	Arg	Glu 510		Pro
	Lys	Glu	Thr 515		Asn	Ala	Ser	Phe 520		Gly	His	Met	Thr 525		Pro	Asn
50	Phe	Lys 530		Glu	Thr	Glu	Ala 535		Glu	Ser	Gly	Leu 540	Glu	Ile	His	Thr
	Val 545	Cys	Ser	Gln	Lys	Glu 550		Ser	Leu	Сув	Pro 555	Asn	Leu	Ile	Asp	Asn 560
	_				565					570			Ala		575	
55				580					585				Lys	590		
			595	_				600					Ile 605			
60		610					615					620	Glu			
	Phe 625	Glu	Ala	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	ser	Gly	Leu	ьeu	640

					645					650				Glu	655	
5				660					665					Сув 670		
			675					680					685	Leu		
10	-	690		_			695					700		Ile		
10	705		_			710					715			Glu Leu		720
		-		_	725					730				Asp	735	
15				740					745					750 Ala		
	Leu		755 Leu	Thr	Pro	Thr		760 Lys	Asp	Val	Leu		765 Asn	Leu	Val	Met
20		770 Ser	Arg	Gly	Lys	Glu 790	775 Ser	Tyr	Lys	Met	Ser 795	780 Asp	Lys	Leu	Lys	Gly 800
	785 Asn	Asn	Tyr	Glu	Ser 805		Val	Glu	Leu	Thr 810		ASN	Ile	Pro	Met 815	
25				820	Val				825	Glu				Asn 830		
			835					840					845	Ser		
		850					855					860		Lys		
30	865					870					875			Asp Val		880
					885					890				His	895	
35				900					905					910 Thr		
	Leu	Tyr	915 Gly	Asp	Thr	Gly	Asp	920 Lys	Gln	Ala	Thr		925 Val	Ser	Ile	Lys
40	-	930 Asp	Leu	Val	Tyr		935 Leu	Ala	Glu	Glu		940 Lys	Asn	Ser	Val	
	945 Gln	His	Ile	Lys	Met 965	950 Thr	Leu	Gly	Gln	Asp 970	955 Leu	Lys	Ser	Asp	Ile 975	960 Ser
45	Leu	Asn	Ile	Asp 980		Ile	Pro		Lys 985	Asn	Asn	Asp		Met 990	Asn	Lys
			995	Leu				Ile 1000	Ser	Asn			1005	5		Ser
		101	)				101	5				1020	)	His		
50	Lys 102	-	Ser	Lys	Met	Phe 1030		Lys	Asp	Ile	Glu 103		Gln	Tyr	Pro	Thr 104
			Ala	Сув		Glu		Val	Asn		Leu		Leu	Asp	Asn 1055	Gln
	Lys	Lys	Leu				Gln	Ser				Val	Ser	Ala	His	
55	Gln	Ser				Val	Ser				Asn	Ser	His	1070   Ile		Pro
	Gln	Met			Ser	Lys	Gln 109			Asn	Ser	Asn 110	His	Asn	Leu	Thr
60	Pro 110	Ser		Lys	Ala	Glu 111	Ile		Glu	Leu	Ser	Thr		Leu	Glu	Glu 112
			Ser	Gln	Phe			Thr	Gln	Phe			Pro	Ser	Tyr	

		1125		1130	1135
		Ser Thr Phe	114	5	Met Thr Ile Leu 1150
5	115	5	1160		His Val Ile Met 1165
	1170		1175	1180	
10	1185	119	0	1195	Lys Asn Asp Cys 120
		1205		1210	Glu Val Gly Phe 1215
		1220	122	5	Val Ser Thr Glu 1230
15	123	5	1240		Glu Asn Ile Ser 1245
	1250		1255	1260	
20	1265	127	0	1275	Asn His Asn Asp 128
	-	1285		1290	Ile Leu Gln Asn 1295
		1300	130	5	Ile Thr Glu Asn 1310
25	131	15	1320		Thr Ala Ala Ser 1325
	1330		1335	1340	
30	1345	135	0	1355	Leu Phe Thr Asp
		1365		1370	Met Lys Glu Gly 1375
		1380	138	5	Phe Leu Glu Val 1390
35	139	95	1400		Asn Lys Glu Gln 1405
	1410		1415	1420	
40	1425	143	30	1435	Ser Val Ala Lys 144
		1445		1450	Lys Pro Glu Glu 1455
		1460	146	5	Asp Ile Arg Lys 1470
45	14	75	1480		Ile Val Lys His 1485
	1490		145	1500	
50	1505	151	10	1515	Lys Glu Pro Thr 152
		1525		1530	Lys Ile Ala Lys 1535
		1540	154	15	Lys Glu Gln Gly 1550
55	15	55	1560		Lys Thr Leu Lys 1565
	1570		1575	1580	
60	1585	159	90	1595	Ser Leu Asn Asn 160
	Asp Lys As	n Leu Val Sei 1605	r Ile Glu Th	r Val Val Pro 1610	Pro Lys Leu Leu 1615

				1620	)				1629	5				Ser 1630	)	
5	Ile	Phe	Leu 1635		Val	Lys	Val	His 1640		Asn	Val	Glu	Lys 164	Glu 5	Thr	Ala
	Lys	Ser 1650		Ala	Thr	Сув	Tyr 1655		Asn	Gln	Ser	Pro 1660		Ser	Val	Ile
	Glu 1669		Ser	Ala	Leu	Ala 1670		Tyr	Thr	Ser	Cys 1675		Arg	Lys	Thr	Ser 168
10	Val	Ser	Gln	Thr	Ser 1689		Leu	Glu	Ala	Lys 1690		Trp	Leu	Arg	Glu 1699	
	Ile	Phe	Asp	Gly 1700		Pro	Glu	Arg	Ile 1709		Thr	Ala	Asp	Tyr 1710		Gly
15	Asn	Tyr	Leu 1715	_	Glu	Asn	Asn	Ser 1720		Ser	Thr	Ile	Ala 172	Glu 5	Asn	Asp
	•	1730	)				1735	5	_		_	1740	)	Asn		
	1745	5				1750	)				1755	5		Asn		176
20	Gly	Tyr	Leu	Ser	Lys 1765		Lys	Leu	Asp	Ser 1770		Ile	Glu	Pro	Val 1775	
	ГÀа	Asn	Val	Glu 1780		Gln	Lys		Thr 1789		Phe	Ser	Lys	Val 1790		Ser
25	Asn	Val	Lys 1795	-	Ala	Asn	Ala	Tyr 1800		Gln	Thr	Val	Asn 1809	Glu 5	Asp	Ile
	-	1810	)				1815	5				1820	)	Asn	_	
	Ala 1829		Ile	Lys	Leu	Ser 1830		Ser	Asn	Ser	Asn 1835		Phe	Glu	Val	Gly 184
30					1845	5				1850	)			Val	1855	5
	Glu	Thr	Ile	Lys 1860		Val	Lys	Asp	Ile 1869		Thr	Asp	Ser	Phe 1870		Lys
35			1875	5				1880	)				1885			
	Ile	Met 1890		Gly	Cys		Glu 1895		Leu	Asp	Asp	Ser 1900		Asp	Ile	Leu
	190	5			_	1910	?				1915	5		His		192
40	Phe	Ala	qaA	Ile	Gln 1925		Glu	Glu	Ile	Leu 1930		His	Asn	Gln	Asn 1935	
		-		1940	) <sup>"</sup>				1945	5				Val 1950	)	
45														His 5		
	Val	Ser 1970		Ala	Asn	Thr	Cys 1975		Ile	Phe	Ser	Thr 1980		Ser	Gly	Lys
	Ser 198		Gln	Val	Ser	Asp 1990		Ser	Leu	Gln	Аын 1995		Arg	Gln	Val	Phe 200
50					200	5				2010	)			Val	2015	;
	_			2020	כ				202	5				Asn 2030	)	
55	Ile	Arg	Thr 2035		Glu	aiH	Leu	Ile 2040		Gln	Lys	Gly	Phe 204!	Ser 5	Tyr	Asn
	Val	Val 205		Ser	Ser	Ala	Phe 205		Gly	Phe	Ser	Thr 2060		Ser	Gly	Lys
	206	5				207	)				2075	5	_	Gly		208
60	Glu	Glu	Phe	Asp	Leu 208		Arg	Thr	Glu	His 209		Leu	His	Tyr	Ser 2095	
	Thr	Ser	Arg	Gln	Asn	Val	Ser	ГХв	Ile	Leu	Pro	Arg	Val	Asp	Lys	Arg

		21	00		2105		211	0
		2115	-	212	:0	_	2125	Ser Lys
5	213	0		2135		214	.0	Ser Glu
	2145		215	0		2155		Gln Gln 216
10			n Leu Val 2165		217	70		2175
		218			2185	-	219	0
		2195		220	0		2205	Thr Asn
15	221	0		2215		222	0	Phe Glu
	2225		223	0		2235	_	Glu Leu 224
20	_	_	Leu Pro 2245		225	0		2255
		226			2265	-	227	0
٥٢		2275	Leu Ile	228	0		2285	•
25	229	0	Phe Asp	2295		230	0	
	2305		Ser Thr	0		2315		232
30			2325		233	0		2335
		234			2345		235	0
35	_	2355	Leu Ser	236	0	_	2365	
33	237	0	Ser Asn	2375		238	0	
	2385		Arg Asn 239	0		2395		240
40	_	_	Val Phe 2405 Gln Cys		241	0	_	2415
		242	_	_	2425		243	0
45		2435	. Glu Ile	244	0		2445	
	245	0	Thr Phe	2455		246	0	
	2465		247		0,0 011	2475		248
50	Ile Thr	Ser Leu	Gln Asn 2485	Ala Arg	Asp Ile 249	-	Met Arg	Ile Lys 2495
		250		•	2505		251	0
		2515	Thr Leu	252	0		2525	
55	253	0	Ser Ala	2535	•	254	0	
	2545		Cys Ile 255	0		2555		256
60			Glu Asp 2565		257	0		2575
	га GIA	Ile Gli 25	ı Leu Ala 30	Asp Gly	Gly Trp 2585	Leu Ile	Pro Ser 259	

Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro 2600 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr 5 2615 2620 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys 2630 2635 Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu 2650 2645 10 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile 2665 2660 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu 2680 2685 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser 15 2700 2695 Ser Asn Lys Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu 2710 2715 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu 2725 2730 20 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile 2740 2745 2750 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu 2760 2765 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg 25 2775 2780 Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro 2790 2795 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly 2805 2810 30 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu 2820 2825 2830 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu 2835 2840 2845 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala 35 2855 2860 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr 2870 2875 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg 2885 2890 40 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala 2900 2905 2910 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala 2920 2925 Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile 45 2935 2940 Gln Leu Glu Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln 2950 2955 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser 2965 2970 2975 50 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro 2980 2985 2990 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile 3000 3005 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn 55 3015 3020 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val 3030 3035 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His 3045 3050 60 Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val 3065 Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala

		3075					3080					3085			
	Pro Phe	)				3095	5				3100	)			
5	Phe Trp 3105		_		3110	)				3115	5				312
	Ala Ala			3125	;				3130	)				3135	5
10	Thr Leu		3140					3145	5				3150	)	
	Gly His	3155	5				3160	)				3165	5		
	Ile Asp	ס		_		3175	5				3180	)			
15	His Ala 3185				3190	)				3195	5				320
	Gly Pro	_		3205	5				3210	)				3215	5
20	Met Ser		3220		_			3225	5				3230	)	
	Cys Met	3235	5				3240	)				3245	5		
	Thr Ser	)				3255	5				3260	)			
25	Cys Lys 3265				3270	)				3275	;				328
	Pro Pro			3285	;				3290	)				3295	i
30	Ala Phe		3300					3305	5				3310	)	
	Lys Lys	3315	5				3320	)				3325	5		
	Asn Glu 3330	0				3335	5				3340	)			
35	Ala Leu 3345				3350	)				3355	5				336
	Gln Phe			3365	5				3370	)				3375	;
40	Glu Asp	_	3380					3389	5				3390	)	
	Glu Gln	Glu 3399		Ser	Gln	Ala	Ser 3400		Glu	Glu	Cys	Glu 3405		Asn	Lys
	Gln Asp		Ile	Thr	Thr	Lys 3415		Tyr	Ile						
45		(2)	INF	ORMA	TION	I FOF	SEÇ	) ID	NO:	12:					
	(:	i) SI	EQUEN	CE C	HAR	ACTE	RISTI	cs:							
50			LENG					oair	3						
		(C)	STRA	NDEI	ONES	3: <b>d</b> c	ouble	2							
	(		MOLEC												
55	(	ix) l	FEATU	RE:											
			) NAM						ace						
60			) ОТН						A2 (0	OMI5)	)				
	(	xi)	SEQUE	ENCE	DES	CRIP	rion	: SE	QID	NO:	12:				

5	GGTGGCG TCTGCTG ACAGATT CTGGAGC	CGC C	TCGGGT ACCGGC	GTC T GCG G	TTTG( TTTT	CGGC TGTC	G GT A GC	GGGT(	CGCC FCCG	GCC	ggga Aaaa	GAA AAG . A AT	GCGT( AACT( G CC	GAGGGG GCACCT	60 120 180 237
10	GGA TCC Gly Ser 5														285
15	AAC AAA Asn Lys 20														333
	TCT TCA Ser Ser														381
20	CAT AAA His Lys	Asn .													429
25	AAA CCA Lys Pro														477
30	CAA GGG Gln Gly 85											_			525
35	AAA TTC Lys Phe 100														573
40	AGT CTT Ser Leu			Lys					_				_		621
10	TGT CCA Cys Pro	Leu :									_	_		_	669
45	TGT ACA Cys Thr														717
50	TTG TTT Leu Phe 165														765
55	ATT TCT Ile Ser 180														813
60	AGT TCT Ser Ser			r Pro									_	_	861
60	AGA AAT Arg Asn														909

		215			220			225		
5		AGC Ser								957
10		ATC Ile								1005
15	 	 AGT Ser								1053
13		TGC Cys								1101
20		GTA Val 295								1149
25		TGT Cys								1197
30		AAG Lys								1245
35		AAA` Lys	Ser							1293
33		GAA Glu								1341
40		CCC Pro 375								1389
45		TTG Leu								1437
50		CAG Gln								1465
==		ATT Ile								1533
55		TTT Phe								1581
60		TCG Ser 455								1629

5			CAG Gln						1677
10			ATA Ile						1725
10			TCT Ser						1773
15			TTT Phe 520						1821
20			TCT Ser						1869
25			TCC Ser						1917
30			ACA Thr						1965
30			AAA Lys						2013
35			TAT Tyr 600						2061
40			TGT ayD						2109
45			GCA Ala						2157
50			TCA Ser						2205
			GGG Gly						2253
55			ACA Thr 680						2301
60			GAA Glu						2349

e	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
5	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
10	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	750	Thr	GAC Asp	Pne	GIN	755	2493
15	Gln	Lys	Ser	Leu	Leu 760	Tyr	Asp	His	Glu	Asn 765	Ala	Ser	ACT Thr	Leu	770	Leu	2541
20	Thr	Pro	Thr	Ser 775	Lys	Asp	Val	Leu	Ser 780	Asn	Leu	Val	ATG Met	785	ser	Arg	2589
25	Gly	Lys	Glu 790	Ser	Tyr	Lys	Met	Ser 795	Asp	Lys	Leu	Lys	GGT Gly 800	Asn	Asn	Tyr	2637
	Glu	Ser 805	Asp	Val	Glu	Leu	Thr 810	Lys	Asn	Ile	Pro	Met 815	GAA Glu	гÀг	Asn	GIII	2685
30	Asp 820	Val	Cys	Ala	Leu	Asn 825	Glu	Asn	Tyr	Lys	Asn 830	Val	GAG Glu	Leu	Leu	835	2733
35	Pro	Glu	Lys	Tyr	Met 840	Arg	Val	Ala	Ser	Pro 845	Ser	Arg	AAG Lys	vaı	850	Pne	2781
40	Asn	Gln	Asn	Thr 855	Asn	Leu	Arg	Val	Ile 860	Gln	Lys	Asn	CAA Gln	865	GIU	THE	2829
45	Thr	Ser	1le 870	Ser	Lys	Ile	Thr	Val 875	Asn	Pro	Asp	Ser	GAA Glu 880	GIU	Leu	Phe	2877
	Ser	Asp 885	Asn	Glu	Asn	Asn	Phe 890	Val	Phe	Gln	Ile	Ala 895		Glu	Arg	Asn	2925
50	Asn 900	Let	ı Ala	. Leu	Gly	905	Thr	. PÀs	Glu	Leu	His 910	GIu	ACA Thr	Asp	Leu	915	2973
55	Сув	va:	l Ası	ı Glu	920	) )	Phe	: Lys	a Asn	925	Thr	Met	GTT Val	Leu	930	GIĀ	3021
60	Ası	Th:	r Gly	935 935	Lys	Glr	n Ala	a Thi	940	val	. Ser	: Ile	. гла	945	Asp	Leu	3069
	GT.	r TA	T GT	r cr	r GC	A GAC	G GAG	AA E	C AA	CAA A	AGT	GT#	AAG	CAG	CAT	ATA .	3117

	Val Tyr	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960	Gln	His	Ile	
5	AAA ATO Lys Met	Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	TAA naA	ATA Ile	3165
10	GAT AAA Asp Lys 980	A ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	GAC Asp	TÀ2 YYY	TGG Trp	GCA Ala	GGA Gly 995	3213
15	CTC TTA	A GGT	Pro	ATT Ile 1000	TCA Ser	AAT Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg L010	ACA Thr	3261
20	GCT TC	r Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser 1020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
20	AAA AT Lys Me	G TTC t Phe 1030	Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr 1040	AGT Ser	TTA Leu	GCT Ala	3357
25	TGT GT Cys Va 104	l Glu	ATT Ile	GTA Val	Asn	ACC Thr 1050	TTG Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
30	AGC AA Ser Ly 1060	G CCT s Pro	CAG Gln	Ser	ATT Ile 1065	Asn	ACT Thr	GTA Val	Ser	GCA Ala 1070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
35	GTA GT Val Va	T GTT l Val	Ser	GAT Asp 1080	Cys	AAA Lys	AAT Asn	Ser	CAT His 1085	ATA Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	TTA Leu	3501
10	TTT TC Phe Se	C AAG	CAG Gln 1095	Asp	TTT Phe	AAT Asn	Ser	AAC Asn 1100	CAT His	AAT Asn	TTA Leu	Thr	CCT Pro 1105	AGC Ser	CAA Gln	3549
40	AAG GC Lys Al	A GAA a Glu 1110	ılle	ACA Thr	GAA Glu	CTT Leu	TCT Ser 1115	Thr	ATA Ile	TTA Leu	Glu	GAA Glu 1120	Ser	GGA Gly	AGT Ser	3597
45	CAG TI Gln Ph 112	e Glu	A TTT	ACT Thr	CAG	TTT Phe	Arg	AAA Lys	CCA Pro	Ser	TAC Tyr 1135	Ile	TTG Leu	CAG Gln	AAG Lys	3645
50	AGT AC Ser Th	CA TT	GAA Glu	GTC Val	CCT Pro	Glu	AAC Asn	CAG Gln	ATG Met	ACT Thr 1150	Ile	TTA Leu	AAG Lys	ACC	ACT Thr 1155	3693
55	TCT GA	AG GAI	A TGC	AGA Arg 1160	J Asp	GCI Ala	GAT Asp	CTI Leu	CAT His 1165	val	ATA Ile	ATG Met	AAT Asn	GCC Ala 1170	Pro	3741
<b></b>	TCG AT	TT GG	r CAC y Glr 1175	ı Val	A GAC l Asp	AGC Sei	AGC Sei	AAC Lys 1180	Glr	TTI Phe	GAZ Glu	A GGT 1 Gly	Thr 1185	· Val	GAA Glu	3789
60	ATT A	AA CG ys Ar	G AAG g Lya	G TT	r GCT e Ala	r GGG a Gly	CTC	TTC Let	AAJ 1 Lys	AA A	GA(	TGT Cys	AAC aak	AAA Lys	A AGT S Ser	3837

PCT/US98/16905 WO 99/09164

GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Jos Thr Glu Ala Leu Gln AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala ACT ARA ACG GAG CAA AAT ATA ARA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT 

Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe

5	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605 4653
	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4033
10	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
15	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
20	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
25	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1555	4893
30	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG  Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu  1560 1565 1570	4941
35	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
40	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
45	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Inn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 . 1635	5133
50	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
55	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
60	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277

		-
	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
10	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
15	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
20	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
30	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
35	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
40	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
50	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
55	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
60	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG	6045

Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val

PCT/US98/16905

WO 99/09164 TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 

TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT Lys Val Phe Val Pro Phe Lys Thr Lys Ser His Phe His Arg Val 

5	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
10	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
15	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu 1le Thr Ser 2470 2475 2480	7677
20	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CCA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
25	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
30	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
35	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
40	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
45	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2595	8013
	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	8061
50	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
55	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640	8157
60	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655	8205

	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675	8253
5	ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT  Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser  2680 2685 2690	8301
10	GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705	8349
15	ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720	8397
20	GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro ro Leu Leu Ala Val 2725 2730 2735	8445
0.5	TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755	8493
25	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
30	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
35	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
40	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
4.5	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835	8733
45	TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Lys Glu Ala 2840 2845 2850	8781
50	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 . 2860 2865	8829
55	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
60	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT	8973

	Asp Gly Ala Glu Leu Tyr Glu Ala Val 2900 2905	Lys Asn Ala Ala Asp Pro Ala 2910 2915
5	TAC CTT GAG GGT TAT TTC AGT GAA GAG Tyr Leu Glu Gly Tyr Phe Ser Glu Glu 2920	CAG TTA AGA GCC TTG AAT AAT 9021 Gln Leu Arg Ala Leu Asn Asn 2925 2930
10	CAC AGG CAA ATG TTG AAT GAT AAG AAA His Arg Gln Met Leu Asn Asp Lys Lys 2935 2940	Gln Ala Gln He Gln Leu Glu
15	ATT AGG AAG ACC ATG GAA TCT GCT GAA Ile Arg Lys Thr Met Glu Ser Ala Glu 2950 2955	CAA AAG GAA CAA GGT TTA TCA 9117 Gln Lys Glu Gln Gly Leu Ser 2960
	AGG GAT GTC ACA ACC GTG TGG AAG TTG Arg Asp Val Thr Thr Val Trp Lys Leu 2965 2970	CGT ATT GTA AGC TAT TCA AAA 9165 Arg Ile Val Ser Tyr Ser Lys 2975
20	AAA GAA AAA GAT TCA GTT ATA CTG AGT Lys Glu Lys Asp Ser Val Ile Leu Ser 2980 2985	ATT TGG CGT CCA TCA TCA GAT 9213 File Trp Arg Pro Ser Ser Asp 2990 2995
25	TTA TAT TCT CTG TTA ACA GAA GGA AAG Leu Tyr Ser Leu Leu Thr Glu Gly Lys 3000	AGA TAC AGA ATT TAT CAT CTT 9261 Arg Tyr Arg Ile Tyr His Leu 3005 3010
30	GCA ACT TCA AAA TCT AAA AGT AAA TCT Ala Thr Ser Lys Ser Lys Ser Lys Ser 3015 3020	Glu Arg Ala Asn Ile Gin Leu
35	GCA GCG ACA AAA AAA ACT CAG TAT CAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln 3030 3035	CAA CTA CCG GTT TCA GAT GAA 9357 Gln Leu Pro Val Ser Asp Glu 3040
	ATT TTA TTT CAG ATT TAC CAG CCA CGG Ile Leu Phe Gln Ile Tyr Gln Pro Arg 3045 3050	G GAG CCC CTT CAC TTC AGC AAA 9405 G Glu Pro Leu His Phe Ser Lys 3055
40	TTT TTA GAT CCA GAC TTT CAG CCA TCT Phe Leu Asp Pro Asp Phe Gln Pro Ser 3060 3065	T TGT TCT GAG GTG GAC CTA ATA 9453 C Cys Ser Glu Val Asp Leu Ile 3070 3075
45	GGA TTT GTC GTT TCT GTT GTG AAA AAA Gly Phe Val Val Ser Val Val Lys Lys 3080	A ACA GGA CTT GCC CCT TTC GTC 9501 S Thr Gly Leu Ala Pro Phe Val 3085 3090
50	TAT TTG TCA GAC GAA TGT TAC AAT TTA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu 3095	1 Leu Ala Ile Lys Phe Trp Ile
55	GAC CTT AAT GAG GAC ATT ATT AAG CCT Asp Leu Asn Glu Asp Ile Ile Lys Pro 3110 3115	T CAT ATG TTA ATT GCT GCA AGC 9597 O His Met Leu Ile Ala Ala Ser 3120
•	AAC CTC CAG TGG CGA CCA GAA TCC AA Asn Leu Gln Trp Arg Pro Glu Ser Ly 3125 3130	A TCA GGC CTT CTT ACT TTA TTT 9645 s Ser Gly Leu Leu Thr Leu Phe 3135
60	GCT GGA GAT TIT TCT GTG TTT TCT GC Ala Gly Asp Phe Ser Val Phe Ser Al	T AGT CCA AAA GAG GGC CAC TTT 9693 a Ser Pro Lys Glu Gly His Phe

	3140		3145		3150	31!	55
5		lu Thr Phe		Lys Asn T		AAT ATT GAC AS Asn Ile Asp I 3170	
10						CTG CAT GCA AN Leu His Ala An 3185	
15					sp Cys Thr S	CCA GGG CCG TA Ser Gly Pro Ty 100	
13		la Gln Ile 1		Thr Gly A		TTG ATG TCT TO Leu Met Ser Se	
20						CTT TGT ATG GO eu Cys Met Al 323	.a
25		rg Lys Ser V		Pro Val S		TG ACT TCA AA let Thr Ser Ly 3250	
30						AC TGC AAA AA sn Cys Lys Ly 3265	
35			sp Phe Leu		eu Pro Leu P	CT CCA CCT GI ro Pro Pro Va 80	
33		ro Ile Cys T		Ser Pro A		AG GCA TTT CA ys Ala Phe Gl	
40						TA AAG AAA AA le Lys Lys Ly 331	s
45		eu Asn Ser E		Thr Pro P		TC AAT GAA AT he Asn Glu Il 3330	
50						TT GCA TTG AT eu Ala Leu Il 3345	
- <b>-</b>			eu Leu Ser		hr Gly Glu L	AA CAA TTT AT ys Gln Phe Il 60	
55		al Ser Glu S		Thr Ala P		CA GAA GAT TA er Glu Asp Ty	
60						AA GAA CAG GA ys Glu Gln Gl 339	.u

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 5 3400 3405 10485 ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415 10 (2) INFORMATION FOR SEQ ID NO:13: (i) SEQUENCE CHARACTERISTICS: 15 (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 20 (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13: 25 Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 10 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 30 40 Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 55 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 35 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 90 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser 105 100 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 120 40 115 125 Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val 135 140 Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val 150 155 Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr 45 170 Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met 185 Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val 50 200 Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp 215 220 Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu 230 235 55 Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr 245 250 Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn 265 Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro 60 275 280 His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu 295 300

	305				Ser	310					315					320
5		_		_	Thr 325					330					335	
				340	Сув				345					350		
			355		Glu			360					365			
10		370			Gln	_	375					380				
	385				Pro Gly	390					395					400
15		-			405 Gln					410					415	
			-	420	Lys				425					430		
20		-	435		Pro			440					445			
		450			Asp	-	455					460				
	465 Ile	Leu	Ala	Val	Lys	470 Gln	Ala	Ile	Ser	Gly	475 Thr	Ser	Pro	Val		480 Ser
25	Ser	Phe	Gln	_	485 Ile	Lys	Lys	Ser		490 Phe	Arg	Ile	Arg		495 Ser	Pro
	Lys	Glu		500 Phe	Asn	Ala	Ser	Phe 520	505 Ser	Gly	His	Met	Thr 525	510 Asp	Pro	Asn
30	Phe	Lys 530	515 Lys	Glu	Thr	Glu	Ala 535		Glu	Ser	Gly	Leu 540		Ile	His	Thr
	Val 545		Ser	Gln	Lys	Glu 550		Ser	Leu	Сув	Pro 555		Leu	Ile	Asp	Asn 560
35	Gly				Ala 565					570					575	
		_		580	Ser				585					590		
			595		Glu			600					605			
40		610			Leu		615					620				
	625				Leu Arg	630					635					640
45				_	645 Ser					650					655	
				660	Ser				665					670		
50	Lys	Glu	675 Ala	Lys	Cys	Asn	Lys	680 Glu	Lys	Leu	Gln		685 Phe	Ile	Thr	Pro
		690 Ala	Asp	Ser	Leu		695 Cys	Leu	Gln	Glu		700 Gln	Cys	Glu	Asn	
55	705 Pro	Lys	Ser	Lys	Lys	710 Val	Ser	Asp	Ile	Lys 730	715 Glu	Glu	Val	Leu	Ala 735	720 Ala
55	Ala	Сув	His	Pro	725 Val	Gln	His	Ser	Lys 745		Glu	Tyr	Ser	Asp 750		Asp
	Phe	Gln	Ser 755		Lys	Ser	Leu	Leu 760		Asp	His	Glu	Asn 765		Ser	Thr
60		770	Leu		Pro		775	Lys				780	Asn			
	Ile	Ser	Arg	Gly	Lys	Glu	Ser	Tyr	Lys	Met	Ser	qaA	ГЛE	Leu	Lys	Gly

	785			790					795					800
	Asn Asn	Tyr G	lu Ser 805		Val	Glu	Leu	Thr 810		Asn	Ile	Pro	Met 815	
5	Lys Asn	8	20	_			825			-	_	830		
	Leu Leu	835				840					845			
10	Val Gln 850				855					860				
	Glu Glu 865			870					875					880
1.5	Glu Leu		885					890					895	
15	Glu Arg	9	00				905					910		
	Asp Leu	915	_			920					925			
20	Leu Tyr 930				935					940				
	Lys Asp 945 Gln His		-	950					955	_				960
25	Leu Asn		965					970					975	
25	Trp Ala	9	80				985					990		
	Phe Arg	995		_		1000	)				1005	i		
30	101 Lys Lys	0			1015	,				1020	)			
	1025 Ser Leu			1030	}				1035	,				104
			1045	<b>,</b>				1050	1				1055	,
35	Lys Lys		060		Gln	Ser	Ile 1065	;				1070	)	
				_							His	םו ד	Thr	Pro
	Gln Ser	Ser V 1075				1080	Cys )				1085	5		
40	Gln Met	Ser V 1075 Leu P 0	he Ser	Lys	Gln 1099	1080 Asp	Cys ) Phe	Asn	Ser	Asn 1100	1085 His	Asn	Leu	Thr
40	Gln Met 109 Pro Ser 1105	Ser V 1075 Leu P O Gln L	he Ser Lys Ala	Lys Glu	Gln 1099 Ile	1080 Asp Thr	Cys ) Phe Glu	Asn Leu	Ser Ser 1115	Asn 1100 Thr	1085 His Ile	Asn Leu	Leu Glu	Thr Glu 112
40	Gln Met 109 Pro Ser 1105 Ser Gly	Ser V 1075 Leu P O Gln L	Phe Ser Lys Ala Gln Phe 1125	Lys Glu 1110 Glu	Gln 1095 Ile ) Phe	1080 Asp Thr	Cys ) Phe Glu Gln	Asn Leu Phe 1130	Ser Ser 1115 Arg	Asn 1100 Thr Lys	1085 His Ile Pro	Asn Leu Ser	Leu Glu Tyr 1135	Thr Glu 112 Ile
40	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln	Ser V 1075 Leu F 0 Gln L Ser G	Phe Ser Lys Ala Sin Phe 1125 Ser Thr	Lys Glu 1110 Glu Phe	Gln 1099 Ile ) Phe Glu	1080 Asp Thr Thr	Cys Phe Glu Gln Pro 1145	Asn Leu Phe 1130 Glu	Ser Ser 1115 Arg	Asn 1100 Thr Lys	1085 His Ile Pro	Asn Leu Ser Thr	Leu Glu Tyr 1135 Ile	Thr Glu 112 Ile Leu
	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1 Thr S	Phe Ser Lys Ala Eln Phe 1125 Ger Thr 1140 Ger Glu	Lys Glu 1110 Glu Fhe Glu	Gln 1099 Ile Phe Glu Cys	1080 Asp Thr Thr Val Arg	Cys Phe Glu Gln Pro 1145 Asp	Asn Leu Phe 1130 Glu Ala	Ser Ser 1115 Arg Asn Asp	Asn 1100 Thr Lys Gln Leu	1085 His Ile Pro Met His 1165	Asn Leu Ser Thr 1150 Val	Leu Glu Tyr 1135 Ile	Thr Glu 112 Ile Leu Met
	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1 Thr S 1155 Pro S	Phe Ser Lys Ala Eln Phe 1125 Ger Thr 1140 Ger Glu Ser Ile	Lys Glu 1110 Glu Phe Glu Gly	Gln 1095 Ile Phe Glu Cys Gln 1175	1080 Asp Thr Thr Val Arg 1160 Val	Cys Phe Glu Gln Pro 1145 Asp	Asn Leu Phe 1130 Glu Ala Ser	Ser Ser 1115 Arg Asn Asp	Asn 1100 Thr Lys Gln Leu Lys 1180	1085 His Ile Pro Met His 1165 Gln	Asn Leu Ser Thr 1150 Val Phe	Leu Glu Tyr 1135 Ile Ile Glu	Thr Glu 112 Ile Leu Met Gly
45	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala 117 Thr Val	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1 Thr S 1155 Pro S 0 Glu I	Phe Ser Lys Ala Eln Phe 1129 Ger Thr 1140 Ger Glu Ger Ile	Clu Clu Fhe Glu Glu Gly Arg	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys	1080 Asp Thr Thr Val Arg 1160 Val	Phe Glu Gln Pro 1145 Asp Asp Ala	Asn Leu Phe 1130 Glu Ala Ser Gly	Ser Ser 1115 Arg Asn Asp Ser Leu 1195	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu	IOSS His Ile Pro Met His IIGS Gln	Asn Leu Ser Thr 1150 Val Phe Asn	Leu Glu Tyr 1135 Ile Ile Glu Asp	Thr Glu 112 Ile Leu Met Gly Cys 120
45	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala 117 Thr Val 1185 Asn Lys	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1155 Pro S 0 Glu I	Phe Ser Lys Ala Eln Phe 1125 Ger Thr 1140 Ger Glu Ger Ile Ile Lys Ala Ser	Glu Glu Fhe Glu Gly Arg 1190 Gly	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys	1080 Asp Thr Thr Val Arg 1160 Val Phe	Phe Glu Gln Pro 1149 Asp Asp Ala	Asn Leu Phe 1130 Glu Ala Set Gly Asp 1210	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu	IOSE His Ile Pro Met His Iles Gln Lys	Asn Leu Ser Thr 1150 Val Phe Asn	Tyr 1135 Ile Ile Glu Asp Gly 1215	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe
45	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala 117 Thr Val 1185 Asn Lys	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1 Thr S 1155 Pro S 0 Glu I	Phe Ser Lys Ala Eln Phe 1125 Ser Thr 1140 Ser Glu Ser Ile Lys Ala Ser 1205 Tyr Ser	Glu Glu Fhe Glu Gly Arg 1190 Gly Arg Arg Ala	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr	Thr Thr Val Arg 1160 Val Phe Leu Gly	Phe Glu Gln Pro 1149 Asp Asp Ala Thr	Asn Leu Phe 1130 Glu Ala Set Gly Asp 1210 Lys	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn	IOSE His Ile Pro Met His Iles Gln Lys Glu Val	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230	Tyr 1135 Ile Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe
<b>4</b> 5	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala 117 Thr Val 1185 Asn Lys Arg Gly Ala Leu	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1155 Pro S 0 Glu I Ser P Phe T 11235	Phe Ser Lys Ala Eln Phe 1125 Ger Thr 1140 Ger Glu Ser Ile Lys Ala Ser 1205 Tyr Ser	Glu Glu Fhe Glu Gly Arg 1190 Gly Arg Arg Arg Arg Sly Ala	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr His	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly Leu 1240	Phe Glu Gln Pro 1149 Asp Asp Ala Thr Thr 1229 Phe	Asn Leu Phe 1130 Glu Ala Set Gly Asp 1210 Lys Ser	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn Asn	IOSE His Ile Pro Met His 1165 Gln Lys Glu Val	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn	Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser
<b>4</b> 5	Gln Met 109 Pro Ser 1105 Ser Gly Leu Gln Lys Thr Asn Ala 117 Thr Val 1185 Asn Lys Arg Gly Ala Leu Glu Glu	Ser V 1075 Leu P 0 Gln L Ser G Lys S 1155 Pro S 0 Glu I Ser F Phe T 11235 Thr S	Phe Ser Lys Ala Eln Phe 1125 Ser Thr 1140 Ser Glu Ser Ile Lys Ala Ser 1205 Tyr Ser	Glu Glu Phe Glu Gly Arg 1190 Gly Arg Arg Cly Ala Val	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr His Lys	1080 Asp Thr Thr Val Arg 1160 Val Fhe Leu Gly Leu 1240 His	Phe Glu Gln Pro 1149 Asp Asp Ala Thr Thr 1229 Phe	Asn Leu Phe 1130 Glu Ala Set Gly Asp 1210 Lys Ser Ile	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn Asn Ile Leu 1260	IOSE His Ile Pro Met His 1165 Gln Lys Glu Val Glu 1245 Ser	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn Ser	Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser Lys

					1285	;				1290	)			Leu	1299	5
5				1300	)				1305	5				Thr 1310	)	
	-	_	1315	i				1320	)				1325			
		1330	)				1335	;				1340	)	Ser		
10	1345	5		_		1350	)				1355	5		Phe		136
					1365	i				1370	)			Lys	1375	5
15				1380	)				1385	5				Leu 1390 Lys	)	
		•	1395	5				1400	)				1405			
		1410	)				1415	5				1420	)			
20	1425	5				1430	)				1435	5		Val		144
					1445	5				1450	)			Pro	1455	<b>i</b>
25				1460	)				1465	5				Ile 1470	)	
		_	1475	5				1480	)				1485			
		1490	)				1495	5				1500	)	Gln		
30	1509	5				1510	)				1515	5		Glu		152
					1525	5				1530	)			Ile	1535	,
35				1540	)				1545	5				Glu 1550	)	
			1555	5				1560	)				1565			
	_	1576	)				1575	5				1580	)	Thr		
40	1589	5				1590	)				1595	5		Leu		160
					1609	5				1610	)			Lys	1615	5
45				1620	)				1625	5				Ser 1630	)	
			163	5				1640	)				1645			
	-	165	0				1 u 5!	5				1660	)	Ser		
50	166	5				1676	0				167	5		ràs		168
	Val	Ser	Gln	Thr	Ser 168		Ļеu	Glu	Ala	Lys 1690		Trp	Leu	Arg	Glu 169	
55				170	0				170	5				Tyr 171	)	
	Asn	Tyr	Leu 171		Glu	Asn	naA	Ser 172		Ser	Thr	Ile	Ala 172	Glu 5	Asn	qaA
	Lys	Asn 173	His		Ser	Glu	Lys 173		Asp	Thr	Tyr	Leu 174		Asn	Ser	Ser
60	Met 174	Ser	Asn	Ser	Tyr	Ser 175	Tyr		Ser	Asp	Glu 175		Tyr	Asn	Asp	Ser 176
	Gly	Tyr	Leu	Ser	Lys			Leu	Asp	Ser			Glu	Pro	Val	Leu

					1765					1770					1775	
	Lys	Asn	Val			Gln	Lys		Thr 1785		Phe	Ser	Lys	Val 1790		Ser
5	Asn	Val	Lys 1795			Asn	Ala		Pro		Thr	Val	Asn 1805	Glu		Ile
	Сув		Glu		Leu	Val	Thr 1815	Ser		Ser	Pro	Cys 1820	Lys	Asn	Lys	Asn
1.0			Ile	Lys	Leu	Ser 1830	Ile		Asn	Ser	Asn 1835	Asn		Glu	Val	Gly 184
10	1825 Pro	Pro	Ala	Phe		Ile		Ser	Gly	Lys 1850	Ile		Сув	Val		His
	Glu	Thr	Ile	Lys 1860			Lys		Ile 1865	Phe		Asp	Ser	Phe 1870	Ser	
15	Val	Ile	Lys 1879	Glu	Asn	Asn	Glu		Lys		Lys	Ile	Cys 1885	Gln		Lys
	Ile		Ala	Gly	Cys	Tyr		Ala		Asp	Asp	Ser 1900	Glu	Asp	Ile	Leu
			Ser	Leu	Asp			Glu	Cys	Ser	Thr	His		His	Lys	Val 192
20	1909 Phe	5 Ala	Asp	Ile			Glu	Glu	Ile				Asn	Gln	Asn	Met
	Ser	Gly	Leu			Val	ser	Lys				Cys	Asp	Val	1935 Ser	
25	Glu	Thr			) Ile	Cys	Lys				Gly	Lys	Leu	1950 His		Ser
	Val	Ser	1959 Ser	5 Ala	Asn	Thr				Phe	Ser	Thr	1969 Ala	Ser	Gly	Lys
	Ser	197 Val	0 Gln	Val	Ser	Asp	197! Ala	5 Ser	Leu	Gln	Asn	1980 Ala		Gln	Val	Phe
30	198	5				1990	) 	T	<b>01</b> 5	1751	1999		Lare	va 1	T.em	200 Phe
					2009	5				2010	)			Val	2015	•
	-			202	0				2025	5				Asn 2030	)	
35			203	5				204	0				204			
		205	n				205	5				206	0	Ser		
40	206	_				237	0				207	5		Gly		208
					208	5				209	0			Tyr	2095	•
				210	٥				210	5				Asp 211	)	
45			211	5				212	0				212			
		213	10				213	5				214	0	Ser		
50	214	15				215	0				215	5		Phe		216
	Asp	Lys			216	5				217	0			Val	217	5
				218	30				218	5				Val 219	U	
55			219	95				220	10				220			
		227	ս Val 10	Cys			221	L5				222	0	Tyr		
60	223	r Gli 25	u Ala			223	30				223	5		Asp		224
	Th	r As	p Se	r Lys	s Lev 224		Se	r His	a Ala	Th:	His	Ser	Leu	ı Phe	Thr 225	Cys 5

				2260	)				226	5				227		
5	Arg	-	2275	5				2280	)				228	5		
		2290	)				229	5				2300	)			
	2305					2310	)				2315	5			Arg	232
10					2325	5				2330	)				Phe 233!	5
	Thr		-	2340	. ັ				234!	5				235	0	
15	Gly	Gln	Glu 2355		Leu	Ser	ГÀЗ	Ser 2360		Leu	Tyr	Glu	His 236		Thr	Leu
	Glu	Lys 2370		Ser	Ser	Asn	Leu 2379		Val	Ser	Gly	His 2380		Phe	Tyr	Gln
	Val 2385		Ala	Thr	Arg	Asn 2390		Lys	Met	Arg	His 2399		Ile	Thr	Thr	Gly 240
20	Arg	Pro	Thr	Lys	Val 2409		Val	Pro	Pro	Phe 2410	_	Thr	Lys	Ser	His 241	
	His	Arg	Val	Glu 2420		Cys	Val	Arg	Asn 2425		Asn	Leu	Glu	Glu 2430		Arg
25	Gln	Lys	Gln 2435		Ile	Asp	Gly	His 2440		Ser	qaA	qaA	Ser 244		Asn	Lys
		Asn 2450	-	Asn	Glu	Ile	His 2455		Phe	Asn	Lys	Asn 2460		Ser	Asn	Gln
	Ala 2465		Ala	Val	Thr	Phe 2470		Lys	Суѕ	Glu	Glu 2475		Pro	Leu	qaA	Leu 248
30	Ile	Thr	Ser	Leu	Gln 2485		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2499	
	Lys	Lys	Gln	Arg 2500		Arg	Val	Phe	Pro 2505		Pro	Gly	Ser	Leu 2510		Leu
35	Ala	Lys	Thr 2515		Thr	Leu	Pro	Arg 2520		Ser	Leu	Lys	Ala 252		Val	Gly
	Gly	Gln 2530		Pro	Ser		Cys 2535		His	Lys	Gln	Leu 2540		Thr	Tyr	Gly
	Val 2545		Lys	His	Cys	Ile 2550		Ile	Asn	Ser	Lys 2555		Ala	Glu	Ser	Phe 256
40	Gln	Phe	His	Thr	Glu 2569	-	Tyr	Phe	Gly	Lys 2570		Ser	Leu	Trp	Thr 2575	_
	Lys	Gly	Ile	Gln 2580		Ala	Asp	Gly	Gly 2585		Leu	Ile	Pro	Ser 2590		Asp
45	Gly	-	Ala 2595	_	ГÀв									Asp		Pro
	Gly	Val 2610		Pro	Lys	Leu	Ile 2615		Arg	Ile	Trp	Val 2620		naA	His	Tyr
	Arg 2625		Ile	Ile	Trp	Lys 2030		Ala	Ala	Met	Glu 2635		Ala	Phe	Pro	Lув 264
50	Glu	Phe	Ala	Asn	Arg 2645	_	Leu	Ser	Pro	Glu 2650		Val	Leu	Leu	Gln 2655	
	Lys	Tyr	Arg	Tyr 2660		Thr	Ģlu	Ile	Asp 2665		Ser	Arg	Arg	Ser 2670		Ile
55	Lys	Lys	Ile 2675	Met		Arg	Asp	Asp 2680		Ala	Ala	Lys	Thr 2689		Val	Leu
	-	Val 2690		qaA	Ile	Ile	Ser 2699		Ser	Ala	Asn	Ile 2700		Glu	Thr	Ser
		Asn		Thr	Ser	Ser 271	Ala		Thr	Gln	Lys 2715	Val		Ile	Ile	Glu 272
60			Asp	Gly	Trp 2725	Tyr		Val	Lys	Ala 2730	Gln		Asp	Pro	Pro 2735	
	Leu	Ala	Val	Leu			Gly	Arg	Leu			Gly	Gln	Lys	Ile	

				274					274	_				2750		
			275	5				2760	כ	Pro	_		2765	5		
5	Glu	Ala 2770		Glu	Ser	Leu	Met 277		Lys	Ile	Ser	Ala 2780		Ser	Thr	Arg
	Pro 2789		Arg	Trp	Tyr	Thr 2790	_	Leu	Gly	Phe	Phe 2795		qaA	Pro	Arg	Pro 280
10	Phe	Pro	Leu	Pro	Leu 2805		Ser	Leu	Phe	Ser 2810		Gly	Gly	Asn	Val 281	_
	Cys	Val	Asp	Val 2820		Ile	Gln	Arg	Ala 2829	Tyr 5	Pro	Ile	Gln	Trp 2830		Gli
	Lys	Thr	Ser 2835		Gly	Leu	Tyr	Ile 2840		Arg	Asn	Glu	Arg 2845		Glu	Gli
15	Lys	Glu 2850		Ala	Lys	Tyr	Val 2859		Ala	Gln	Gln	Lys 2860	_	Leu	Glu	Ala
	Leu 286		Thr	Lys	Ile	Gln 2870		Glu	Phe	Glu	Glu 2875		Glu	Glu	Asn	Th: 288
20	Thr	Lys	Pro	Tyr	Leu 2885		Ser	Arg	Ala	Leu 2890		Arg	Gln	Gln	Val 2899	
	Ala	Leu	Gln	Asp 2900	_	Ala	Glu	Leu	Tyr 2905	Glu 5	Ala	Val	Lys	Asn 2910		Ala
	Asp	Pro	Ala 2915	_	Leu	Glu		Tyr 2920		Ser	Glu		Gln 2925		Arg	Ala
25	Leu	Asn 2930		His	Arg		Met 2935		Asn	Asp	_	Lys 2940		Ala	Gln	Ile
	2945	5			_	2950	)				2955			_		296
30					2965	;				Trp 2970					2975	5
				2980	)				2985					2990	)	
			2995	;				3000	}	Glu			3005			
35	_	3010	)				3015	;		Ser		3020		_		
	3025	5				3030	)	_			3035					304
40					3045	i				Gln 3050					3055	i
			-	3060	)	_			3065				_	3070		
45			3075	;				3080	)	Val	-		3085	_		
13		3090	)				3095	;		Tyr Ile		3100				
	3105	5		_		3110	١				3115					312
50					3125					3130 Phe					3135	
				3140	)				3145					3150	_	
55			3155	i				3160		Asn			3165			
		3170	)				3175	;		Pro		3180				
	3185	5		_		3190	)				3195	_	_	-		320
60					3205					3210 Tyr				_	3215	
				3220		-,-			3225	_				3230		

	cys me	323	_	AIG	пув	ser	324		1111	PIO	vaı	324		GIII	Mec	
5	Thr Se	50		_	-	325	5	_			326	ס <u> </u>		_		
	Cys Ly 3265	s Lys	Arg	Arg	Ala 3270		Asp	Phe	Leu	Ser 327		Leu	Pro	Leu	Pro 328	
	Pro Pr	o Val	Ser	Pro 3289		Сув	Thr	Phe	Val 3290	Ser		Ala	Ala	Gln 3295	Lys	
10	Ala Ph	e Gln	Pro 3300	Pro		Ser	Сув	Gly 3309	Thi		Tyr	Glu	Thr 3310	Pro		
	Lys Ly	s Lys 331:	Glu		Asn	Ser	Pro 3320	${\tt Gln}$		Thr	Pro		Lys		Phe	
15	Asn Gl	ı Ile		Leu	Leu		Ser		Ser	Ile				Glu	Leu	
13	Ala Le		Asn	Thr			-	Leu	Ser	_			Gly	Glu	_	
	3345 Gln Ph	e Ile	Ser	Val	3350 Ser		Ser	Thr	Arg	3355 Thr		Pro	Thr	Ser	336 Ser	
20	Glu Asj	o Tyr	Leu	3365 Arg		Lys	Arq	Arg	3370 Cys		Thr	Ser	Leu	3375 Ile		
	Glu Gli		3380	)				3385	5				3390	)	-	
		339	5				3400	)		014	C, S	3405	_	ASII	272	
25	Gln Ası 34:		116	1111	1111	3415		IYI	116							
		(2)	INF	ORMA	TION	FOR	SEC	QI Q	NO:1	4:						
		(i) SI	QUEN	ICE C	HARA	CTER	ISTI	CS:								
30			LENG				-	s								
			STRA					!								
35		\_,														
33			NAM													
			LOC			ITAM	ON:	2F p	rime	r						
40	(	xi) S	EQUE	NCE	DESC	RIPT	ION:	SEQ	ID :	NO:1	4:					
	TGAGTTI	TAC C	CTCAG	TCAC	Α											20
		(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:1	6:						
45	(	i) SE	QUEN	CE C	HARA	CTER	ISTI	CS:								
			LENG					s								
50		(C)	STRA	NDED	NESS	: si	ngle									
50		(ח)	TOPO	LOGY	: 11	near										
	(	xi) S	EQUE	NCE	DESC	Ript	ION:	SEQ	ID 1	NO:1	6 :					
55	CAGGAAA	.CAG C	TATG	ACCC	T GT	GACG	TACT	GGG	TTTT	TAG	С					41
		(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:1	7 :						
60	(	i) SE	_													
50		(B)	LENG TYPE	: nu	clei	c ac	id									
		(C)	STRA	NDED	NESS	: si	nale									

	(D) TOPOLOGY: linear	
5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3FII primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
10	GATCTTTAAC TGTTCTGGGT CACA	24
	(2) INFORMATION FOR SEQ ID NO:18:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) DPOLOGY: linear	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3RII primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:	
	CCCAGCATGA CACAATTAAT GA	22
30	(2) INFORMATION FOR SEQ ID NO:19:	
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 44 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 4F/M 13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
45	TGTAAAACGA CGGCCAGTAG AATGCAAATT TATAATCCAG AGTA	44
	(2) INFORMATION FOR SEQ ID NO:20:	
50	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 22 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
55	(A) NAME/KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 4R-1A primer	

ATCAGATTCA TCTTTATAGA AC

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	(2) INFORMATION FOR SEQ ID NO:21:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
10		
	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 5+6F/M13F primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:	
	TGTAAAACGA CGGCCAGTTG TGTTGGCATT TTAAACATCA	40
20	(2) INFORMATION FOR SEQ ID NO:22:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 38 base pairs	
0.5	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
25	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
30	(B) LOCATION:	
	(D) OTHER INFORMATION: 5+6R/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
35	CAGGAAACAG CTATGACCCA GGGCAAAGGT ATAACGCT	38
	(2) INFORMATION FOR SEQ ID NO:23:	
	(i) SEQUENCE CHARACTERISTICS:	
40	(A) LENGTH: 38 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(b) Toronoor. Illien	
45	(T) WANT /VEV	
	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 7F/M13F primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:	
		38
	TGTAAAACGA CGGCCAGTTA AGTGAAATAA AGAGTGAA	
	(2) INFORMATION FOR SEQ ID NO:24:	
55	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 36 base pairs	
	(B) TYPE: nucleic acid	
<b>C</b> C	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
60	(D) IOFOHOGI. IIIIOMI	

5	<ul> <li>(A) NAME/KBY:</li> <li>(B) LOCATION:</li> <li>(D) OTHER INFORMATION: 7R/M13R primer</li> <li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:</li> </ul>	
	CAGGAAACAG CTATGACCAG AAGTATTAGA GATGAC	36
10	(2) INFORMATION FOR SEQ ID NO:25:	
15	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8F/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:	
25	TGTAAAACGA CGGCCAGTGC CATATCTTAC CACCTTGTGA	40
	(2) INFORMATION FOR SEQ ID NO:26:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(ix) FEATURE:	
40	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 8FIA primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:	
	TTGCATTCTA GTGATAATAT AC	22
45	(2) INFORMATION FOR SEQ ID NO:27:	
50	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 19 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
55	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 8RIA primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	1:
60	AATTGTTAGC AATTTCAAC  (2) INFORMATION FOR SEQ ID NO:28:	
	\@/ AATA VAMEEEE EE: E EE: A TE E EE:	

5	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9F/M13F primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:	40
	TGTAAAACGA CGGCCAGTTG GACCTAGGTT GATTGCAGAT	40
	(2) INFORMATION FOR SEQ ID NO:29:	
20	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
25	(2) 10102001. 22-002	
	<ul><li>(A) NAMB/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9R/M13R primer</li></ul>	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:	
	CAGGAAACAG CTATGACCTA AACTGAGATC ACGGGTGACA	40
35	(2) INFORMATION FOR SEQ ID NO:30:	
40	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 24 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
45	(A) NAME/KEY: (B) LOCATION: (D) OTHFR INFORMATION: 10AF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:	24
50	GAATAATATA AATTATATGG CTTA	27
	(2) INFORMATION FOR SEQ ID NO:31:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 37 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 10AR/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:	
5	CAGGAAACAG CTATGACCCC TAGTCTTGCT AGTTCTT	37
	(2) INFORMATION FOR SEQ ID NO:32:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 42 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
15		
0.0	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10BF/M13F primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:	
	TGTAAAACGA CGGCCAGTAR CTGAAGTGGA ACCAAATGAT AC	42
25	(2) INFORMATION FOR SEQ ID NO:33:	
30	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 44 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
	(D) TOPOLOGI: IIIlear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10BR/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	
40	CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA	44
	(2) INFORMATION FOR SEQ ID NO:34:	
45	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
50	(ix) FEATURE:	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10CF/M13F primer</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:	
	TGTAAAACGA CGGCCAGTCA GCATCTTGAA TCTCATACAG	40
60	(2) INFORMATION FOR SEQ ID NO:35:	
	(i) SEQUENCE CHARACTERISTICS:	

5	<ul><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10CRII primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:	
	AGACAGAGGT ACCTGAATC	19
15	(2) INFORMATION FOR SEQ ID NO:36:	
20	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
25	(A) NAME/KEY:	
23	(B) LOCATION: (D) OTHER INFORMATION: 11AF-M13 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:	
30	TGTAAAACGA CGGCCAGTTG GTACTTTAAT TTTGTCACTT	40
	(2) INFORMATION FOR SEQ ID NO:37:	
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 37 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	\ <del>-</del> /	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11AR-M13 primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:</li></ul>	
	CAGGAAACAG CTATGACCTG CAGGCATGAC AGAGAAT	37
50	(2) INFORMATION FOR SEQ ID NO:38:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11BF primer</li></ul>	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:	
	AAGAAGCAAA ATGTAATAAG GA	22
5	(2) INFORMATION FOR SEQ ID NO:39:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11BR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:	
20	CATTTAAAGC ACATACATCT TG	22
	(2) INFORMATION FOR SEQ ID NO:40:	
25	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 21 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single	
30	(D) TOPOLOGY: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11CF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:</li></ul>	
	TCTAGAGGCA AAGAATCATA C	21
40	(2) INFORMATION FOR SEQ ID NO:41:	
45	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
50	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11CR primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:	
66	CAAGATTATT CCTTTCATTA GC	22
55	(2) INFORMATION FOR SEQ ID NO:42:	
60	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 22 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single	

	(D) TOPOLOGY: linear	
5	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11DF primer	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:	
	AACCAAAACA CAAATCTAAG AG	22
	(2) INFORMATION FOR SEQ ID NO:43:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11DR primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:	
	GTCATTTTTA TATGCTGCTT TAC	23
30	(2) INFORMATION FOR SEQ ID NO:44:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11EF primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:	
45	GGTTTTATAT GGAGACACAG G	21
40	(2) INFORMATION FOR SEQ ID NO:45:	
50	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 23 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
55	(A) NAME/KBY:	
	(B) LOCATION: (D) OTHER INFORMATION: 11ER primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:	
	GTATTTACAA TTTCAACACA AGC	23

	(2) INFORMATION FOR SEQ ID NO:46:	
5	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 20 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
10		
	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 11FF primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:	
	(X1) SEQUENCE DESCRIPTION: SEQ 1D NO. 40.	
	ATCACAGTTT TGGAGGTAGC	20
20	(2) INFORMATION FOR SEQ ID NO:47:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 21 base pairs	
	(B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: single	
23	(D) TOPOLOGY: linear	
	(6, 2000_000)	
	(A) NAME/KEY:	
30	(B) LOCATION:	
30	(D) OTHER INFORMATION: 11FR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:	
35	CTGACTTCCT GATTCTTCTA A	21
	(2) INFORMATION FOR SEQ ID NO:48:	
	(2) INFORMATION FOR SEQ ID NO. 40.	
	(i) SEQUENCE CHARACTERISTICS:	
40	(A) LENGTH: 21 base pairs	
-	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
4 =		
45	(A) NAMF/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 11GF primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:	
	CTCAGATGTT ATTTTCCAAG C .	2
	(2) INFORMATION FOR SEQ ID NO:49:	
55	(1) OPENINGE CHADACTURITETICS.	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li></ul>	
	(A) LENGTH: 21 base pairs (B) TYPE: nucleic acid	
	(B) TYPE: NUCLEIC ACIU (C) STRANDEDNESS: single	
60	(C) STRANDEDNESS: BINGLE (D) TOPOLOGY: linear	
OU	(D) IOEOTHORI TIME	

	(A) NAME/KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 11GR primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:	
		21
	CTGTTAAATA ACCAGAAGCA C	
10	(2) INFORMATION FOR SEQ ID NO:50:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 18 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
20	(B) LOCATION:	
	(D) OTHER INFORMATION: 11HF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:	
25	AGGTAGACAG CAGCAAGC	18
	(2) INFORMATION FOR SEQ ID NO:51:	
	(i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 22 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
35	(ix) FEATURE:	
	(A) NAME/KEY: None	
	(B) LOCATION:	
40	(D) OTHER INFORMATION: 11HR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:	
	GTAATATCAG TTGGCATTTA TT	22
45	(2) INFORMATION FOR SEQ ID NO:52:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 21 base pairs	
50	(B) TYPE: nucleic acid (C) STRANDEDMESS: single	
50	(D) TOPOLOGY: linear	
	<b>(2)</b>	
	(A) NAME/KEY:	
55	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 111F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:	
60	TGCAGAGGTA CATCCAATAA G	21
	(2) INFORMATION FOR SEQ ID NO:53:	
	(4) AMA WAMER - WEST TOTAL TOT	

5	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 21 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
10	<ul><li>(A) NAME/KBY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11IR primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:	
	GATCAGTAAA TAGCAAGTCC G	21
	(2) INFORMATION FOR SEQ ID NO:54:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
25	(D) TOPOLOGY: linear	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11JF primer</li></ul>	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:	
	TACTGAAAAT GAAGATAACA AAT	23
35	(2) INFORMATION FOR SEQ ID NO:55:	
55		
40	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: !!JR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:	
50	ATTTTGTTCT TTCTTATGTC AG	22
	(2) INFORMATION FOR SEQ ID NO:56:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 35 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60		
	(A) NAME/KEY: (B) LOCATION:	

	(D) OTHER INFORMATION: 11KF-M13 primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:	
5	TGTAAAACGA CGGCCAGTCT ACTAAAACGG AGCAA	35
	(2) INFORMATION FOR SEQ ID NO:57:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 35 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
15	,-,	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11KR-M13 primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:	
	CAGGAAACAG CTATGACCGT ATGAAAACCC AACAG	35
25	(2) INFORMATION FOR SEQ ID NO:58:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	(D) TOPOLOGY: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11LF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:	
40	CACAAAATAC TGAAAGAAAG TG	22
	(2) INFORMATION FOR SEQ ID NO:59:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50		
	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11LR primer	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:	
	GGCACCACAG TCTCAATAG	19
60	(2) INFORMATION FOR SEQ ID NO:60:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs	

	(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
5		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MF primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:	
	GCAAAGACCC TAAAGTACAG	20
15	(2) INFORMATION FOR SEQ ID NO:61:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	(D) TOPOLOGY: linear	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:	
30	CATCAAATAT TCCTTCTCTA AG	22
	(2) INFORMATION FOR SEQ ID NO:62:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	(A) NAME, KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 11NF-M13 primer	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:	
	TGTAAAACGA CGGCCAGTGA AAATTCAGCC TTAGC	35
50	(2) INFOPMATION FOR SEQ ID NO:63:	
	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 35 base pairs  (B) TYPE: nucleic acid  (C) SERAMBERNESS: single	
55	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11NR-M13 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:	

CAGGAAACAG CTATGACCAT CAGAATGGTA GGAAT	35
(2) INFORMATION FOR SEQ ID NO:64:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 22 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(b) Topologi. Timeat	
<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 110F primer</li></ul>	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:	
GTACTATAGC TGAAAATGAC AA	22
(2) INFORMATION FOR SEQ ID NO:65:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pair</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 110R primer</li></ul>	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:	
ACCACTGGCT ATCCTAAATG	20
(2) INFORMATION FOR SEQ ID NO:66:	
(A) LENGTH: 20 base pairs (B) TYPE: nucleic acid	
(D) TOPOLOGY: linear	
(A) NAME/KEY: (B) LOCATION:	
(D) OTHER INFORMATION: 11PF primer	
	20
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
	(2) INFORMATION FOR SEQ ID NO:64:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 110F primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:  GTACTATAGC TGAAAATGAC AA  (2) INFORMATION FOR SEQ ID NO:65:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pair (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 110R primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:  ACCACTGGCT ATCCTAAATG  (2) INFORMATION FOR SEQ ID NO:66: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11PF primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:  TGAAGATATT TGCGTTGAGG  (2) INFORMATION FOR SEQ ID NO:67: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: SINGLE (D) TOPOLOGY: linear

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11PR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:	
10	GTCAGCAAAA ACCTTATGTG	20
	(2) INFORMATION FOR SEQ ID NO:68:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11QF primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:	
	ACGAAAATTA TGGCAGGTTG T	21
2.0	(2) INFORMATION FOR SEQ ID NO:69:	
30	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 21 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
35	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
40	(B) LOCATION:	
40	(D) OTHER INFORMATION: 11QR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:	
45	CTTGTCTTGC GTTTTGTAAT G	21
	(2) INFORMATION FOR SEQ ID NO:70:	
	(i) SEQUENCE CHARACTERISTICS:	
50	(A) LENGTH: 20 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
55	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 11RF primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:	
	GCTTCATAAG TCAGTCTCAT	20

	(2) INFORMATION FOR SEQ ID NO:71:	
5	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11RR primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:	
	TCAAATTCCT CTAACACTCC	20
20	(2) INFORMATION FOR SEQ ID NO:72:	
	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 35 base pairs  (B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11SF-M13 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:	
35	TGTAAAACGA CGGCCAGTTA CAGCAAGTGG AAAGC	35
	(2) INFORMATION FOR SEQ ID NO:73:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 37 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11SR-M13 primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:	
50	CAGGAAACAG CTATGACCAA GTTTCAGTTT TACCAAT	37
	(2) INFORMATION FOR SEQ ID NO:74:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
60	(D) TOPOLOGY: linear	

178

(A) NAME/KEY:

	(B) LOCATION: (D) OTHER INFORMATION: 11TF primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:	
	GTTCTTCAGA AAATAATCAC TC	22
10	(2) INFORMATION FOR SEQ ID NO:75:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
15	(D) TOPOLOGY: linear	
20	(A) NAME/KEY:  (C) LOCATION:  (D) OTHER INFORMATION: 11TR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:	
25	TGTAAAAGA GAATGTGTGG C	21
23	(2) INFORMATION FOR SEQ ID NO:76:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11UF-M13 primer	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:	
40	TGTAAAACGA CGGCCAGTAC TTTTTCTGAT GTTCCTGTG	39
	(2) INFORMATION FOR SEQ ID NO:77:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
50	(D) TOPOLOGY: _inear	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11UR-M13 primer</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:	
	CAGGAAACAG CTATGACCTA AAAATAGTGA TTGGCAACA	39
60	(2) INFORMATION FOR SEQ ID NO:78:	
	(1) SPONDENCE CHARACTERISTICS.	

5	<ul><li>(A) LENGTH: 42 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12F/M13F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:</li></ul>	
15	TGTAAAACGA CGGCCAGTAG TGGTGTTTTA AAGTGGTCAA AA  (2) INFORMATION FOR SEQ ID NO:79:	42
20	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
25	<ul><li>(A) NAME/KBY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12R/M13R primer</li></ul>	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:  CAGGAAACAG CTATGACCGG ATCCACCTGA GGTCAGAATA  (2) INFORMATION FOR SEQ ID NO:80:	40
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 21 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	(0)	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 13-2F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:</li></ul>	
	TAACATTTAA GCATCCGTTA C	21
50	(2) INFORMATION FOR SEQ ID NO:81:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 28 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 13-2R primer	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:	
_	AAACGAGACT TTTCTCATAC TGTATTAG	28
5	(2) INFORMATION FOR SEQ ID NO:82:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 14F primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:	
20	ACCATGTAGC AAATGAGGGT CT	22
	(2) INFORMATION FOR SEQ ID NO:83:	
25	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
30	(b) Topology: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 14AR primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:</li></ul>	
	GCTTTTGTCT GTTTTCCTCC AA	22
40	(2) INFORMATION FOR SEQ ID NO:84:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	<ul><li>(A) NAME, KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 15-2F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:	
55	CCAGGGGTTG TGCTTTTAA A	21
	(2) INFORMATION FOR SEQ ID NO:85:	
60	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	

(C) STRANDEDNESS: single

PCT/US98/16905 WO 99/09164

	(D) TOPOLOGY: linear	
5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 15FUT/M13-R primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:	
10	CAGGAAACAG CTATGACCAC TCTGTCATAA AAGCCATC	38
	(2) INFORMATION FOR SEQ ID NO:86:	
15	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 24 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
20		
25	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 16AF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:	24
	TTTGGTTTGT TATAATTGTT TTTA	24
30	(2) INFORMATION FOR SEQ ID NO:87:	
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 16AR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:	
45	CCAACTTTTT AGTTCGAGAG	20
	(2) INFORMATION FOR SEQ ID NO:88:	
50	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 19 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 17F primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:	
	THE STATE OF THE S	19

TTCAGTATCA TCCTATGTG

	(2) INFORMATION FOR SEQ ID NO:89:	
5	(i) SEQUENCE CHARACTERISTICS:	
J	(A) LENGTH: 20 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
10		
	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 17AR primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:	
	(XI) SEQUENCE PERSONNELLE	
	AGAAACCTTA ACCCATACTG	20
	TOP OFFI TO NO. 90.	
20	(2) INFORMATION FOR SEQ ID NO:90:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 39 base pairs	
	(B) TYPE: nucleic acid	
25	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
30	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 18FUT/M13-AF primer</li></ul>	
	(D) OTHER INFORMATION. ISTOT/1115 in promo-	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:	
35	TGTAAAACGA CGGCCAGTGA ATTCTAGAGT CACACTTCC	39
35		
	(2) INFORMATION FOR SEQ ID NO:91:	
	(i) SEQUENCE CHARACTERISTICS:	
40	(A) LENGTH: 38 base pairs	
10	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
45		
13	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 18R/M13R primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:	
54		38
	CAGGAAACAG CTATGACCTT TAACTGAATC AATGACTG	3.
	(2) INFORMATION FOR SEQ ID NO:92:	
55		
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 41 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
60	(D) TOPOLOGY: linear	

	(A) NAME/KEY:	
	(B) LOCATION:	
_	(D) OTHER INFORMATION: 19F/M13F primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:	
	TGTAAAACGA CGGCCAGTAA GTGAATATTT TTAAGGCAGT T	41
10	(2) INFORMATION FOR SEQ ID NO:93:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 39 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
0.0	(A) NAME/KEY:	
20	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 19FUT/M13-R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:	
25	CAGGAAACAG CTATGACCAA GAGACCGAAA CTCCATCTC	39
	(2) INFORMATION FOR SEQ ID NO:94:	
	(i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 38 base pairs	
	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	(D) TOPOLOGY: linear	
35		
	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 20F/M13F primer	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:	2.0
	TGTAAAACGA CGGCCAGTCA CTGTGCCTGG CCTGATAC	38
45	(2) INFORMATION FOR SEQ ID NO:95:	
43	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 39 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRATEDNESS: single	
50	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
	(B) LOCATION:	
55	(D) OTHER INFORMATION: 20R/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:	
60	CAGGAAACAG CTATGACCAT GTTAAATTCA AAGTCTCTA	39
	(2) INFORMATION FOR SEQ ID NO:96:	

5	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 39 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 21F/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:	
15	TGTAAAACGA CGGCCAGTGG GTGTTTTATG CTTGGTTCT	39
	(2) INFORMATION FOR SEQ ID NO:97:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
25	(A) NAME/KEY:	
	(A) NAME/REI: (B) LOCATION: (D) OTHER INFORMATION: 21R/M13R primer	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:	
	CAGGAAACAG CTATGACCCA TTTCAACATA TTCCTTCCTG	40
35	<ul><li>(2) INFORMATION FOR SEQ ID NO:98:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li></ul>	
40	(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
45	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 22F-1A primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:	19
50	AACCACACCC TTAAGATGA	25
	(2) INFORMATION FOR SEQ ID NO:99:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 22R-1A primer</li></ul>	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:	
5	GCATTAGTAG TGGATTTTGC	20
	(2) INFORMATION FOR SEQ ID NO:100:	
10	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 16 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 23FII primer</li></ul>	
20	(x1) SEQUENCE DESCRIPTION: SEQ ID NO:100:	
	TCACTTCCAT TGCATC	16
	(2) INFORMATION FOR SEQ ID NO:101:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 17 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
30	(D) TOPOLOGY: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 23RII primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:</li></ul>	
	TGCCAACTGG TAGCTCC	17
40	(2) INFORMATION FOR SEQ ID NO:102:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24 2F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:	
55	TACAGTTAGC AGCGACAAAA	20
	(2) INFORMATION FOR SEQ ID NO:103:	
60	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	

	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	-
5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24R/M13R primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:	
	CAGGAAACAG CTATGACCAT TTGCCAACTG GTAGCTCC	38
	(2) INFORMATION FOR SEQ ID NO:104:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) TOPOLOGY: linear	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25F-7/23 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:	
	GCTTTCGCCA AATTCAGCTA	20
30	(2) INFORMATION FOR SEQ ID NO:105:	
35	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25R-7/23 primer</li></ul>	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105: TACCAAAATG TGTGGTGATG	20
	(2) INFORMATION FOR SEQ IL NO:106:	
50	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 20 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
55	(a) Assault (VIIV.	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 26-2F primer</li></ul>	
60	(xi) SEGUENCE DESCRIPTION: SEQ ID NO:106:	

	and the second s	PCT/US98/16905
	WO 99/09164	-
	AATCACTGAT ACTGGTTTTG	20
_	(2) INFORMATION FOR SEQ ID NO:107:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li></ul>	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
10	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
15	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 26-2R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:	
0.0	TATACTTACA GGAGCCACAT	20
20	(2) INFORMATION FOR SEQ ID NO:108:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 18 base pairs</li></ul>	
25	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
30	(A) NAME/KEY:	
	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 27AF-1A primer</li></ul>	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:	
35	CTGTGTGTAA TATTTGCG	18
	(2) INFORMATION FOR SEQ ID NO:109:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li></ul>	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
45		
	(A) NAME/KEY: (B) LOCATION:	
50	(D) OTHER INFORMATION: 27AR/M13R primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:	
	CAGGAAACAG CTATGACGGC AAGTTCTTCG TCAGCTATTG	40
55	(2) INFORMATION FOR SEQ ID NO:110:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li></ul>	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
60	(D) TOPOLOGY: linear	

	(A) NAME/KEY:	
	(B) LOCATION:	
5	(D) OTHER INFORMATION: 27BF/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:	
10	TGTAAAACGA CGGCCAGTGA ATTCTCCTCA GATGACTCCA	40
10	(2) INFORMATION FOR SEQ ID NO:111:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li></ul>	
1 5	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
20	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 27BR/M13R primer	
2.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:	
25	CAGGAAACAG CTATGACCTC TTTGCTCATT GTGCAACA	38
	<u> </u>	

## **WE CLAIM:**

5 1. A genomic DNA containing a BRCA2 gene,

wherein the first twelve nucleotides beginning exon 5 are 5'-

TCCTGTTGTTCT-3' as set forth in SEQ. ID. NO: 1,

wherein nucleotides numbers 5782-5790 are GTTTGTGTT as set forth in SEQ. ID. NO: 4, and

wherein the last 20 nucleotides ending exon 15 are 5'CTGCGTGTTCTCATAAACAG-3' as set forth in SEQ. ID. NO: 2 and the first 20
nucleotides beginning exon 16 are 5'-CTGTATACGTATGGCGTTTC-3' as set forth in SEQ. ID. NO: 3.

15 2. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 20 2457 T 2908 G 3199 A 3624 A 4035 T 25 7470 A 9079 G.

3. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

30 1093 A 1342 C 1593 A 2457 T 35 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

4. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 1593 A 2457 T 2908 G 3199 A 3624 A 4035 C 7470 A 9079 G.

5. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 C 1342 A 20 1593 A 2457 C 2908 G 3199 G 3624 G 25 4035 T 7470 G 9079 G.

6. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 1593 A 35 2457 T 2908 G 3199 A 3624 G 4035 T 40 7470 G 9079 G.

7. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 C 1342 C 1593 G 2457 C 50 2908 A 3199 G

3624 A 4035 T 7470 A 5 9079 A.

8. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

10 2024 C 4553 C 4815 G 5841 T 5972 C.

20

9. A DNA comprising a BRCA2 coding sequence, wherein nucleotide numbers 643-666 are

CTTAGTGAAAGTCCTGTTGTTCTA and

wherein nucleotides numbers 5782-5790 are GTTTGTGTT.

10. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 2457 T 2908 G 3199 A 30 3624 A 4035 T 7470 A 9079 G.

11. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

1093 A
1342 C
40 1593 A
2457 T
2908 G
3199 A
3624 A
45 4035 T
7470 A
9079 G
as set forth in SEQ. ID. NO: 4.

12. The DNA according to claim 9 wherein the coding sequence nucleutides are as follows:

PCT/US98/16905 WO 99/09164

```
1093 A
          1342 C
          1593 A
 5
          2457 T
          2908 G
          3199 A
          3624 A
          4035 C
10
          7470 A
          9079 G
     as set forth in SEQ. ID. NO: 6.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 15 13. as follows:

```
1093 C
          1342 A
          1593 A
20
          2457 C
          2908 G
          3199 G
          3624 G
          4035 T
25
          7470 G
          9079 G
     as set forth in SEQ. ID. NO: 8.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 30 14. as follows:

```
1093 A
          1342 C
          1593 A
35
          2457 T
          2908 G
          3199 A
          3624 G
          4035 T
40
          7470 G
          9079 G
     as set forth in SEQ. ID. NO: 10.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 15. 45 as follows:

```
1093 C
          1342 C
          1593 G
50
          2457 C
```

2908 A 3199 G 3624 A 4035 T 5 7470 A 9079 A as set forth in SEQ. ID. NO: 12.

The DNA according to claim 9 wherein the coding sequence nucleotides are 16. 10 as follows:

2024 C 4553 C 4815 G 15 5841 T 5972 C.

A BRCA2 protein having the following amino acids at the following peptide 17. numbers: 20

289 asparagine histidine 372 894 valine 991 asparagine 25 1852 valine 1853 cysteine 1854 valine 2951 alanine

- as set forth in SEQ. ID. NO: 5. 30
  - The BRCA2 protein having the following amino acids at the following peptide 18. numbers:
- 289 asparagine 35 asparagine 372 599 serine valine 894 991 asparagine 2951 alanine. 40
  - The BRCA2 protein having the following amino acids at the following peptide 19. numbers:
- 289 histidine 45 372 histidine valine 894 991 asparatic acid 2951 alanine
- as set forth in SEQ. ID. NO: 9. 50

20. The BRCA2 protein having the following amino acids at the following peptide numbers:

5 289 histidine

372 asparagine

894 isoleucine

991 aspartic acid

2951 threonine

- as set forth in SEQ. ID. NO: 13.
  - 21. The BRCA2 protein according to claims 17-20 having the following amino acids at the following peptide numbers:

15 59? serine

30

1442 serine

1915 threonine.

- 22. A haplotype of BRCA2 coding sequence (BRCA2<sup>omi 1</sup>) as set forth in SEQ. ID.
- NO: 4 or a sequence complementary thereto.
  - 23. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>orni</sup> as set forth in SEQ. ID. NO: 5.
- 24. A haplotype of BRCA2 coding sequence (BRCA2<sup>omi 2</sup>) as set forth in SEQ. ID.
   NO: 6 or a sequence complementary thereto.
  - 25. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>omi</sup>
    <sup>2</sup> as set forth in SEQ. ID. NO: 7.
  - 26. A haplotype of BRCA2 coding sequence (BRCA2<sup>omi 3</sup>) as set forth in SEQ. ID. NO: 8 or a sequence complementary thereto.
- 27. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup>
  3 as set forth in SEQ. ID. NO: 9.

28. A haplotype of BRCA2 coding sequence (BRCA2<sup>oml 4</sup>) as set forth in SEQ. ID. NO: 10 or a sequence complementary thereto.

- <sup>4</sup> as set forth in SEQ. ID. NO: 11.
  - 30. A haplotype of BRCA2 coding sequence (BRCA2<sup>omi 5</sup>) as set forth in SEQ. ID. NO: 12 or a sequence complementary thereto.
  - 31. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>omi</sup> as set forth in SEQ. ID. NO: 13.
- 32. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;
  - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
  - (d) comparing the sequence of said amplified DNA fragment to a
     BRCA2<sup>(orni)</sup> DNA sequence selecting from the group consisting of SEQ.
     ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
     NO: 12, and their respective complementary sequences;
  - (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
    - (i) AAT and CAT at position 1093,
    - (ii) CAT and AAT at position 1342,
    - (iii) TCA and TCG at position 1593,

25

10

5

10

15

25

- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 33. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2
     coding sequence;
  - (b) sequencing said amplified DNA fragment;
  - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
    - (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(orni)</sup> DNA sequence selecting from the group consisting of SEQ.
       ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

(e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:

- (i) AAT and CAT at position 1093,
- (ii) CAT and AAT at position 1342,
- (iii) TCA and TCG at position 1593,
- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence; wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:
- (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively,
  - (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,

25

5

10

15

20

at position 2457, CAT and CAC occur at frequencies from (iv) about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) about 85-95%, and from about 5-15%, respectively, 5 at position 3199, AAC and GAC occur at frequencies (vi) from about 75-85%, and from about 15-25%, respectively. at position 36<sup>4</sup>, AAA and AAG occur at frequencies (vii) from about 75-85%, and from about 15-25%, 10 respectively, at position 4035, GTT and GTC occur at frequencies from (viii) about 85-95%, and from about 5-15%, respectively, at position 7470, TCA and TCG occur at frequencies from (ix) about 75-85%, and from about 15-25%, respectively, and 15 at position 9079, GCC and ACC occur at frequencies (x) from about 85-95%, and from about 5-15%, respectively.

- 34. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;
- 25 (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
  - (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ.

ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

- determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093,

5

10

15

20

- (ii) CAT and AAT at position 1342,
- (iii) TCA and TCG at position 1593,
- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 35. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;

(c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;

- (d) comparing the sequence of said amplified DNA fragment to a
   BRCA2<sup>(oml)</sup> DNA sequence selected from the group consisting of: SEQ.
   ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
   NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of: SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035,
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence, wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:

20

5

10

PCT/US98/16905

WO 99/09164

at position 1093, AAT and CAT occur at frequencies from (i) about 75-85%, and from about 15-25%, respectively, at position 1342, CAT and AAT occur at frequencies from (ii) about 35-45%, and from about 55-65%, respectively, 5 at position 1593, TCA and TCG occur at frequencies from (iii) about 85-95%, and from about 5-15%, respectively, at position 2457, CAT and CAC occur at frequencies from (iv) about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) 10 about 85-95%, and from about 5-15%, respectively, at position 3199, AAC and GAC occur at frequencies (vi) from about 75-85%, and from about 15-25%, respectively, at position 3624, AAA and AAG occur at frequencies (vii) 15 from about 75-85%, and from about 15-25%, respectively, at position 4035, GTT and GTC occur at frequencies from (viii) about 85-95%, and from about 5-15%, respectively, at position 7470, TCA and TCG occur at frequencies from (ix) 20 about 75-85%, and from about 15-25%, respectively, and at position 9079, GCC and ACC occur at frequencies (x)

36. A method according to any of the claims 32-35 wherein the said amplifying is performed by annealing at least one oligonucleotide primer to said DNA fragment and extending the oligonucleotide primer by an agent for polymerization.

from about 25-95%, and from about 5-15%, respectively.

37. A method according to claim 36 wherein said oligonucleotide primer is directly or indirectly labeled with a radioactive label, a fluorescent label, a bioluminescent label, a chemiluminescent label, a metal chelator, or an enzyme label.

5

38. A BRCA2 coding sequence according to claims 32, wherein the codon pairs occur at the following frequencies:

10

(i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,

(ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively,

- (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,

(v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively,

(vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,

20

15

(vii) at position 3624, AAA and AAG occur at frequencies from about 75-85%, and from about 15-25%, respectively,

25

- (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (ix) at position 7470, TCA and TCG occur at frequencies from about 75-85%, and from about 15-25%, respectively, and
- (x) at position 9079, GCC and ACC occur at frequencies from about 85-95%, and from about 5-15%, respectively.

39. An oligonucleotide primer capable of hybridizing to a sample of BRCA2 gene, or its respective complementary sequences selected from the group consisting of SEQ. ID. NO: 14, 19, 22, 23, 25, 26, 29-76, 83, 85-88, 90, 91, 97, 98, 101, and 104-107.

- 40. A chip array having "n" elements for performing allele specific sequencebased techniques comprising a solid phase chip and oligonucleotides having "n" different nucleotide sequences,
- wherein "n" is an interger greater than or equal to ten,

15

20

30

wherein said oligonucleotides are bound to said solid phase chip in a manner which permits said oligonucleotides to effectively hybridize to complementary oligonucleotides or polynucleotides,

wherein oligonucleotides having different nucleotide sequence are bound to said solid phase chip at different locations so that a particular location on said solid phase chip exclusively binds oligonucleotides having a specific nucleotide sequence, and

wherein at least ten oligonucleotides are capable of specifically hybridizing to the BRCA2 DNA having the sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12 or their respective complementary sequences, at least one oligonucleotide being capable of specifically hybridizing at each of the nucleotide positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, 9079, or complementary thereto.

- 25 41. A method of performing gene therapy on a patient, comprising:
  - a) contacting cancer cells *in vivo* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their respective complementary sequences
    - b) allowing the vector to enter the cancer cells, and
    - c) measuring a reduction in tumor growth.
  - 42. The method according to claim 41 wherein said cancer cells have a mutation in the BRCA2 gene.

43. The method according to claim 41 wherein said patient has a mutation in the BRCA2 gene of non-cancer cells.

5

- 44. A method of performing gene therapy on a patient or a sample, comprising:
- a) contacting cells *in vivo* or *in vitro* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO:
- 10 10, SEQ. ID. NO: 12, or their respective complementary sequences, and
  - b) allowing the vector to enter the cells,
     wherein said patient has a reduced susceptibility for developing a cancer associated with a mutation in the BRCA2 gene.
- 15 45. A method according to claim 44 wherein said cells include healthy breast, ovarian or pancreatic tissues.
  - 46. A method according to claim 44 wherein a patient has an inherited mutation in the BRCA2 gene.

20

- 47. A method of treating a patient suspected of having a tumor, comprising:
- a) administering to a patient an effective amount of BRCA2 tumor growth inhibitor having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof;
  - b) allowing the patient's cells to take up the protein, and
  - c) measuring a reduction in tumor growth.
- 48. The method according to claim 47 wherein said tumor is a breast cancer, an ovarian cancer or a pancreatic cancer.
  - 49. The method according to claim 47 wherein said patient has an inherited mutation in the BRCA2 gene.

50. A method of preventing the formation or growth of a tumor, comprising:

a) adminstering to a patient an effective amount of BRCA2 tumor growth inhibiting protein having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof; and

- b) allowing the patient cells to take up the protein.
- 51. The method according to claim 31 wherein the protein is administered parenternally, by buccal adsorption or inhalation.
  - 52. A cloning vector comprising:

- (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof; and
- (b) one or more suitable regulatory sequences to induce replication and/or integration in a host cell.
- 53. An expression vector comprising a DNA sequence as set forth in SEQ. ID.
   NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any
   fragments thereof operatively linked to one or more promoter sequences capable of directing expression of said sequence in a host cell.
  - 54. A host cell transformed with the vector according to claim 52 or 53.
- 25 55. A BRCA2 polypeptide which is selected from the group consisting of:

  (a) a fragment of BRCA2 protein sequence as set forth in SEQ. ID. NO: 5,

  SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO:13;

  (b) an amino acid sequence which is substantially homologous to the BRCA2 protein sequence as set forth in SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID.

  NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO: 13;
  - (c) a molecule which has similar function to the BRCA2 protein; and (d) a fusion protein of (a), (b), or (c).

56. An anti-BRCA2 antibody wherein a molecule according to claims 17-21, 23, 25, 27, 29, 31, or 55 is used as an immunogen.

- 5 57. A diagnostic reagent comprising a molecule selected from the group consisting of:
  - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO:
  - 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their complementary sequences;
  - (b) a nucleic acid fragment of (a) comprising at least 10 nucleotide in length;
- 10 (c) a sequence which hybridizes to (a) or (b);
  - (d) a polypeptide according to claim 17-21, 23, 25, 27, 29, 31, or 55; and
  - (e) an antibody which specifically binds to the polypeptide of (d).
- 58. A pharmaceutical composition comprising a molecule according to any one of the claims 17-21, 23, 25, 27, 29, 31, 55 in a pharmaceutically acceptable carrier.
  - 59. A pharmaceutical composition comprising a molecule according claim 56 in a pharmaceutically acceptable carrier.
- 60. A pharmaceutical composition comprising a molecule according to claim 57 in a pharmaceutically acceptable carrier.

# Figure 1A

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA <u>ATĞ</u>CCTATTGGATCCAAAGAGAGGCCAACATTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

gggattttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG AAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

t cact grant at tig tact gttt cag GAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACAGTGĂAAACŤAAĂATGĞATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

agggatttgctttgtTTATTTTAGTCCTGTTGTTCTACAATGTACACATGTAACACCACAAA GAGATAAGTCAGgtatgattaaaaacaatgctttttattctt

ttaacaattitcccctttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTG **AAGgtaaatatt** 

TCTGAAAGTCTAGGAGCTGAGGTGGATCCTGÄTATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtaataata

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGAŤTŤATCGCTTCTGTGACAGACAGTGAAAACACAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct

ttaatgtgcttctgttttatactttaacagGATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATA GCTGCAAAGACCACATTGGAAAGTCAATGCCAAATGTCCTAGAAGATGAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

# Figure 1B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAAGAAAGATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGACAGCATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTCTGTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACTTACATTTGCAAATGCTGATTCAGgta cctctgtct

ttigtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA PAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 1C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGGAAACACTCAGATTAAAGA^GATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAACATCAAAAAGTATCTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAAAAGCAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAACT:CTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT CTGGTATTGAGCCAGTATTGAAGAATGTTGAAGATCAAAAAAACACTAGTTTTTCCA AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGC ATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTT: GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT. AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 1D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCC~AGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG aagtgttcatttttacctttcgtgttgccaatca

eaeacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATCAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttattata

aatatgtaatataaaattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT Exon 13 TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

ccccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAGAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT GAAAAATGAGACACTTGATTACTACAGGCAGACCAAACCAAAGTCTTTGTTCCACC TTTTAAAACTAAATCaCATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGACAAAACATTGATGGACATGGCTCTGATGATAGTAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

tttttgctaagtatttattctttgatag. TTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC **AAGTTCCCTCTGCgtgtccccataaacaggtatgtgt** 

tuttettuttgtgtgtgtgtttattttgtgtagGTGTTCTCATAAACAGCTGTATACGTATGGCGTTTCTAA ACATTGCATAAAAATTAACAGCAAAAATGCAGAGTCTTTCAGTTTCACACTGAAGA TTATTTTGGTAAGGAAAGTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTG GATGGCTCATACCCTCCAATGATGGAAAGGCTGGAAAAGAAGAAGAATTTTATAGgtactct atgcaaaaagattgtgtgttaacttttatg

# Figure 1E

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGGTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

ttttgttttcacttttagATATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGC AĂŤĞĂĂGAGAGGAAĞĂAAAGGAAGCAGCAAAATATGTGGAGGCCCAACAAAAGA GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtzzzzit agttatatggtacacattgttatttc

agittagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TÄACÄÄGACAGCAAGTTCGTĞCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagaggagtaagaggacatataatgag

tittattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAA TCACAGGCAAATGTTGAATGATAAGAAACAAGCTCAGATCCAGTTGGAAATTAGGA AGGCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAAC CGTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAGAAAAAGATTCAGgtaagtatgt aaatgctttgttttta

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

# Figure 1F

## Exon 24

gaatttttgttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGAAAAAAAACAGgtaatgcacaatatagttaatttttttat tgattcttttaaaaaaacattgtct

## Exon 25

taacattcttttctttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

## Exon 26

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

## Exon 27

tacgttttcattttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAAAAAAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA** 

# Figure 2A

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA Exon 2 <u>ATĞ</u>ČCTAŤŤGGAŤCČAAĂGAGAGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

gggattttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG AAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

tcactgaattattgtactgtttcagGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACAGTGĂAAACTAAĂATGĞATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

agggatttgctttgttttattttagTCCTGTTGTTCTACAATGTACACATGTAACACCACAAAGAG ATAAGTCAGgtatgattaaaaacaatgctttttattctt

ttaacaattttcccctttttttacccccagTGGTATGTGGGGAGTTTGTTTCATACACCAAAGTTTGTG **AAGgtaaatatt** 

taatgatcagggcatttctataaaaaataaactattitcttcctcccagGGTCGTCAGACACACAAACATATT TCTGAAAGTCTAGGAGCTGAGGTGGATCCTGATATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtastaata

tittatcttacagTCAGA AATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGAŤTŤATCGCTTCTGTGACAGACAGTGAAAACACAAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct.

ttaatgtgcttctgttttatactttaacagGATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATA GCTGCAAAGACCACATTGGAAAGTCAATGCCAAATGTCCTAGAAGATGAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

PCT/US98/16905 WO 99/09164

# Figure 2B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAAGAAAGAATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGACAGCATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTC, GTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACITACATTTGCAAATGCTGATTCAGGt acctctgtct

## Exon 11

titgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAAATGATCCAAAAAGCAAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAA I CGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 2C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGCTTTTAT TCTGCTCATGGCACAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGAGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAACAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAAACTTCTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT TAAAAGTAACGAACATTCAGACCAGCTCACAAGAGAAAAAAATACTGCTATACGTA CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 2D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG aadtottcatttttacctttcgtgttgccaatca

## Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATÃATAGAAAATČAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Gotaaaattagctttttatttata

## Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

## Exon 14

coccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTCAAGAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT TTTTAAAACTAAATCACATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGAAAAGCAAAACATTGATGGACATGGCTCTGATGATAGTAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

## Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC AAGTTCCCTCTGCGTGTTCT.CATAAACAGgtatgtgt

## Exon 16

titttcttttttgtgtgtgtttattttgtgtagCTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTA ACAGCĂĂĂĂTGCĂĞAĞTCTTTTCAGTTTCACACTGAAGATTATTTTGGTAAGGAAA GTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTGGATGGCTCATACCCTCC AATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactctatgcaaaaagattgtgtgttaactttt atg

# Figure 2E

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGĞTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

ttttgttttcactttttagATATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

taaatcaatatatttattaatttgtccagATTTCTGCTAACAGTACTCGGCCTGCTCGCTGGTATAC CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGCAĂŤĞĂĂĞAĞAĞAAĞĂAAAĞĞAAĞCAĞCAAAATATĞTĞĞAĞĞCCCAACAAAAĞÂ GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtaaaatt acttatatggtacacattgttatttc

agttagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TĂACĂĂGACAGCAAGŤTCGTĞCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagagagtaagaggacatataatgag

## Exon 22

ttttattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAATCACAGGCAAATGTTGAATGATAAGAAALAAGCTCAGATCCAGTTGGAAATTAGGA AGGCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAACC GTĞTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAGAAAAAGATTCAGgtaagtatgta aatqctttgttttta

## Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGAGATACAGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

## PCT/US98/16905

# Figure 2F

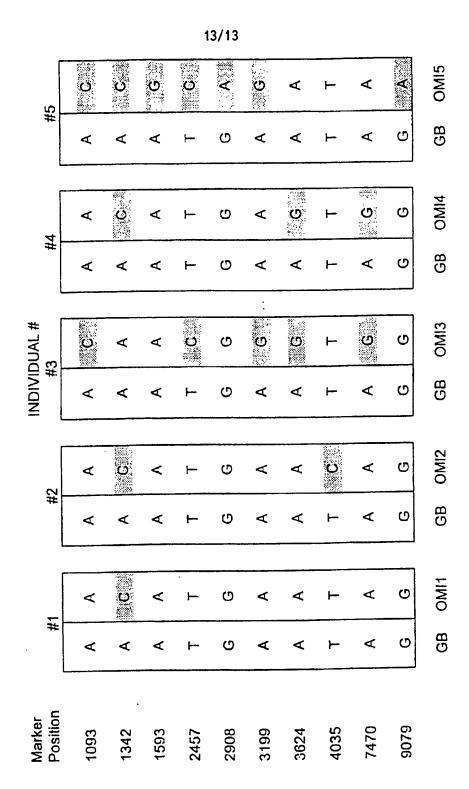
gaatttttgttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGAAAAAAACAGgtaatgcacaatatagttaatttttttat tgattcttttaaaaaacattgtct

taacattcttttcttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

tacgttttcatttttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAAAAAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA** 

# FIGURE 3



# INTERNATIONAL SEARCH REPORT

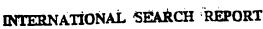
Inte onel Application No PCT/US 98/16905

		<u> </u>	CT/US 98/16905
A. CLASSIF IPC 6	C12N15/12 C07K14/47 C12Q1/68 A61K38/17	A61K48/00	C07K16/18
According to	international Patent Classification (IPC) or to both national classification	on and IPC	
B. FIELDS 8	SEARCHED		
IPC 6	sumentation searched (classification system followed by classification C12N C07K C12Q A61K		d in the fields searched
	on searched other than minimum documentation to the extent that suc		
Electronic da	ata base consulted during the International search (name of data base	and, where practical, se	narch terms used)
C. DOCUME	NTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relev	vant passages	Relevant to claim No.
X	GB 2 307 477 A (CANCER RES CAMPAIO TECHNOLOGY LTD; DUKE UNIVERSITY) 28 May 1997 see figure 7, underlined sequences sheets 20/52, 27/52 and 31/52		1,9, 55-60
X	TAVTIGIAN S V ET AL: "The complete gene and mutations in chromosome 13q-linked kindreds" NATURE GENETICS, vol. 3, no. 12, March 1996, pages XP002076942 cited in the application see table 1C		1-16, 32-39
	-	/	
X Fur	ther documents are listed in the continuation of box C.	X Patent family m	embers are ilsted in annex.
"A" docum consi "E" earlier filling "L" docum which ctatl "O" docur other	nent defining the general state of the art which is not dered to be of particular relevance document but published on or after the international date lent which may throw doubts on priority claim(s) or is cited to establish the publication date of another on or other special reason (as specified) nent reterring to an oral disclosure, use, exhibition or means the published prior to the international filing date but	or priority date and cited to understand invention "X" document of particul cannot be consider involve an inventive "Y" document of particul cannot be consider document is combined in the art.	aned after the international filing date not in conflict with the application but the principle or theory underlying the ar relevance; the claimed invention ed novel or cannot be considered to estep when the document is taken alone ar relevance; the claimed invention are relevance; the claimed invention ed to involve an inventive step when the ned with one or more other such documation being obvious to a person skilled of the same patent family
·	than the priority date claimed  actual completion of the international search		he international search report
	27 January 1999	02/02/19	999
Name and	i mailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk	Authorized officer	
	Tel. (+31-70) 340-2040, Tx. 31 651 epo ni, Fax: (+31-70) 340-3016	Cupido,	M

# INTERNATIONAL SEARCH REPORT

Inte onel Application No PCT/US 98/16905

		PC1/03 98/10903		
	Ition) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevant to claim No.		
Category *	Citation of document, with indication, where appropriate, of the relevant passages			
X	WO 97 22689 A (MYRIAD GENETICS INC ;UNIV PENNSYLVANIA (US); HSC RESEARCH & DEV LI) 26 June 1997 see SEQ ID NO:1 especially nucleotides 643-666 on page 90 and nucleotides 5782-5790 on page 98.	9,55-60		
P,X	5782-5790 on page 98.  WO 97 30108 A (VANDERBILT UNIVERSITY; UNIVERSITY OF WASHINGTON) 21 August 1997 see SEQ ID NO:3 nucleotides 623-634, 5751-5759 and 7794-7834, for claim 9 see also nucleotides 616-634.	1,9		



.....mational application No.

PCT/US 98/16905

# Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet) This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons: 1. X Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Remark: Although claims 41-43 and 46-51 and partially claims 44 and 45 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the vector. because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a). Claims Nos.: 3. Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet) This International Searching Authority found multiple inventions in this international application, as follows: As all required additional search fees were timely paid by the applicant, this International Search Report covers all As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee. As only some of the required additional search fees were timely paid by the applicant, this International Search Report 3. covers only those claims for which fees were paid, specifically claims Nos.: No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: The additional search fees were accompanied by the applicant's protest. Remark on Protest No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

information on patent family members

Inter	nal	Application No
PCT/	/US	98/16905

Patent document cited in search report		Publication date		atent family nember(s)	Publication date
GB 2307477	A	28-05-1997	AU CA EP WO	7635096 A 2238010 A 0858467 A 9719110 A	11-06-1997 29-05-1997 19-08-1998 29-05-1997
WO 9722689	Α	26-06-1997	US AU CA EP NO	5837492 A 1461597 A 2239733 A 0785216 A 982785 A	17-11-1998 14-07-1997 26-06-1997 23-07-1997 17-08-1998
WO 9730108	 А	21-08-1997	AU	2065397 A	02-09-1997